



FEATURES  
 source  
 Delagen, Inc. (US)  
 Location/Qualifiers  
 1..3106  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

## ORIGIN

Query Match 100.0%; Score 3106; DB 6; Length 3106;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGTGAGAGGCTGCCCCGAGAGGACCGCTCTGAGACCGGAGATCGGACCGCAAAA 60  
 1 CATGTGAGAGGCTGCCCCGAGAGGACCGCTCTGAGACCGGAGATCGGACCGCAAAA 60  
 61 CCATGGGTAGCAATCGGGGCGCGCAAGGCGGAGGAGGCTCTCAGAGCTTCGGCGCGGAC 120  
 61 CCATGGGTAGCAATCGGGGCGCGCAAGGCGGAGGAGGCTCTCAGAGCTTCGGCGCGGAC 120  
 121 TCAAGTACAACTCCCGGCTAGAGAACTGAATGCTTTGAGAGGAGTGTGAGTTCCTGC 180  
 121 TCAAGTACAACTCCCGGCTAGAGAACTGAATGCTTTGAGAGGAGTGTGAGTTCCTGC 180  
 181 CTGGGAACATGCGCAAGAGTGGAGAGGAGGCGCCGAGCGCTGGAGTGTGAGTGG 240  
 181 CTGGGAACATGCGCAAGAGTGGAGAGGAGGCGCCGAGCGCTGGAGTGTGAGTGG 240  
 241 CAGTGTGCTGAGCTTCTCTTGTCTCTCCCTCATGAGTGGTGGCTGGTGGACATTC 300  
 241 CAGTGTGCTGAGCTTCTCTTGTCTCTCCCTCATGAGTGGTGGCTGGTGGACATTC 300  
 241 CAGTGTGCTGAGCTTCTCTTGTCTCTCCCTCATGAGTGGTGGCTGGTGGACATTC 300  
 301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGCGCATCTGAGATCACAATAGA 360  
 301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGCGCATCTGAGATCACAATAGA 360  
 361 TCTTTCTGAGAGGCTAGAGAACTCACTCCAGAGATTATAGCTTGGCCAGCGAGG 420  
 361 TCTTTCTGAGAGGCTAGAGAACTCACTCCAGAGATTATAGCTTGGCCAGCGAGG 420  
 421 TGAAGAGAGGCGCTGAGAGCTCAATGAATCCCTGTCTGGGTCCTTACCAAGAA 480  
 421 TGAAGAGAGGCGCTGAGAGCTCAATGAATCCCTGTCTGGGTCCTTACCAAGAA 480  
 481 AGTGGGCTGTAACTGCTTCACTGAGGAGAGTGTATGCTTACTGCTGAGAGTTC 540  
 481 AGTGGGCTGTAACTGCTTCACTGAGGAGAGTGTATGCTTACTGCTGAGAGTTC 540  
 541 GCATCCCCCACTGGGCGAGAGAGTGAATCGGCGCATGAGGCTGTGAGCGAGTGTAA 600  
 541 GCATCCCCCACTGGGCGAGAGAGTGAATCGGCGCATGAGGCTGTGAGCGAGTGTAA 600  
 601 CATTGSCACCCCGAGACGAGGCACTGAATCTTGTGCTAACTCTGTGTGAGCTTTC 660  
 601 CATTGSCACCCCGAGACGAGGCACTGAATCTTGTGCTAACTCTGTGTGAGCTTTC 660  
 661 CATGCGCACCCCGAGACGAGGCACTGAATCTTGTGCTAACTCTGTGTGAGCTTTC 660  
 661 CATGCGCACCCCGAGAGTGTGAGAGAGTCAAGAGCAAGAGCTGTGAGTGTGAGT 720  
 661 CATGCGCACCCCGAGAGTGTGAGAGAGTCAAGAGCAAGAGCTGTGAGTGTGAGT 720  
 721 CCCATGTGAGAGCAGTGAACAGCTTCACTAACCCTGAGCTTCCCAACAGTCCCTAAC 780  
 721 CCCATGTGAGAGCAGTGAACAGCTTCACTAACCCTGAGCTTCCCAACAGTCCCTAAC 780  
 781 CGCATGCGCGCTGCGAGTGGTCTCTGCGGGGAGCGCGACTCTGTGTGAGCTTCACT 840  
 781 CGCATGCGCGCTGCGAGTGGTCTCTGCGGGGAGCGCGACTCTGTGTGAGCTTCACT 840  
 841 TCCGAGAGCTTTGATGTGCTCTCTCTGTGAGAGCATGTGAGAGCTGTGTGAGTGTG 900  
 841 TCCGAGAGCTTTGATGTGCTCTCTCTGTGAGAGCATGTGAGAGCTGTGTGAGTGTG 900  
 901 ATAGGCTGAGGCCCATGGAACCCCAAGCTGTGTGAGGCTGTGTGAGCACTTCTCACT 960

DB 901 ATAGGCTGAGGCCCATGGAACCCCAAGCTGTGTGAGGCTGTGTGAGCACTTCTCACT 960  
 961 CCTACCACTGATCTTCTCTCTCTCCCAAGAGCTCTTCTGTCAGAGCTATACCAATA 1020  
 961 CCTACCACTGATCTTCTCTCTCTCCCAAGAGCTCTTCTGTCAGAGCTATACCAATA 1020  
 1021 CTGACCGGCGCATCTCTGCTTGAAGGCACTTCTCTTCAAGTGTGCCCAAGATGAGCACT 1080  
 1021 CTGACCGGCGCATCTCTGCTTGAAGGCACTTCTCTTCAAGTGTGCCCAAGATGAGCACT 1080  
 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGACATTTGAGAGGCTCTACTATCCAGGCCACT 1140  
 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGACATTTGAGAGGCTCTACTATCCAGGCCACT 1140  
 1141 ACCCGGCCCAATCACTGAGCAATGGAATTAAGAGGCTCCCAACCGGAGCGTGAAG 1200  
 1141 ACCCGGCCCAATCACTGAGCAATGGAATTAAGAGGCTCCCAACCGGAGCGTGAAG 1200  
 1201 TGGGCTTCAAACTCTTCTATCTGTGAGACCCCAAGTACCAAGTGGCTCTGTGACCAAG 1260  
 1201 TGGGCTTCAAACTCTTCTATCTGTGAGACCCCAAGTACCAAGTGGCTCTGTGACCAAG 1260  
 1261 ACTATGTGAGATCAACGGGGAAGATGCTGCGGTGAGAGGCTCCCAATTTGTGTGAGCA 1320  
 1261 ACTATGTGAGATCAACGGGGAAGATGCTGCGGTGAGAGGCTCCCAATTTGTGTGAGCA 1320  
 1321 GCAACACACCAAGATATAGTACAGTCACTTCCATCTGATCACTGATACAGGACACCGG 1380  
 1321 GCAACACACCAAGATATAGTACAGTCACTTCCATCTGATCACTGATACAGGACACCGG 1380  
 1381 TCTAGCTGAGTACTCTCTCTCAAGCTCAACAGACCGGTGCCAGGATGTTCAATGTGCA 1440  
 1381 TCTAGCTGAGTACTCTCTCTCAAGCTCAACAGACCGGTGCCAGGATGTTCAATGTGCA 1440  
 1441 AGACTGACGCTGATCCGAAGAGAACTGGGCTGGACAGCTGGCCGAGATT 1500  
 1441 AGACTGACGCTGATCCGAAGAGAACTGGGCTGGACAGCTGGCCGAGATT 1500  
 1501 AGACTGACGCTGATCCGAAGAGAACTGGGCTGGACAGCTGGCCGAGATT 1500  
 1501 AGACTGACGCTGATCCGAAGAGAACTGGGCTGGACAGCTGGCCGAGATT 1500  
 1561 ATAGTGAATGAGCGTTACTGCGGATGCAATGCAACCCACAGTTGACGTCAAAAACAGT 1560  
 1561 ATAGTGAATGAGCGTTACTGCGGATGCAATGCAACCCACAGTTGACGTCAAAAACAGT 1560  
 1561 TCTGCAAGCCCTCTTCTGGGCTGTGTGACAGTGTCAACGACTGTGGGAGCGGAAGTGA 1620  
 1561 TCTGCAAGCCCTCTTCTGGGCTGTGTGACAGTGTCAACGACTGTGGGAGCGGAAGTGA 1620  
 1621 AGAAGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1621 AGAAGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1681 AGAGCGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1681 AGAGCGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1681 AGAGCGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1681 AGAGCGGCTGCACTGTCTCTCTGAGAGTTCAGAGTTCAGATGGGAAAGTGTCCCTC 1680  
 1741 ACAGGCTGAATGTGCTCTCTTGAACCAATAATACCTACGCTGCCAAATGAGCTCTGTG 1800  
 1741 ACAGGCTGAATGTGCTCTCTTGAACCAATAATACCTACGCTGCCAAATGAGCTCTGTG 1800  
 1741 ACAGGCTGAATGTGCTCTCTTGAACCAATAATACCTACGCTGCCAAATGAGCTCTGTG 1800  
 1801 TGAAGCAAGGCAACCTTGAAGTGAATGGGAAGAGCGGACTGTAAGAGAGTGTCCGATGAG 1860  
 1801 TGAAGCAAGGCAACCTTGAAGTGAATGGGAAGAGCGGACTGTAAGAGAGTGTCCGATGAG 1860  
 1861 TGAAGCAAGGCAACCTTGAAGTGAATGGGAAGAGCGGACTGTAAGAGAGTGTCCGATGAG 1860  
 1861 TGAAGCAAGGCAACCTTGAAGTGAATGGGAAGAGCGGACTGTAAGAGAGTGTCCGATGAG 1860  
 1861 TGAAGCAAGGCAACCTTGAAGTGAATGGGAAGAGCGGACTGTAAGAGAGTGTCCGATGAG 1860  
 1921 ATGCGGAGCGAGGCGAGTGGCTCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 1980  
 1921 ATGCGGAGCGAGGCGAGTGGCTCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 1980  
 1921 ATGCGGAGCGAGGCGAGTGGCTCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 1980  
 1981 TGTGTGGGCTCTGCTCATCTCTCTGAGTGGTGTCTGAGCTCATGCTTCTCAAG 2040

```

Db      1981  TGTGTGGGGCGCTGCTCATCTCTCTGACTGGCTGTCTCTGCAAGCTCATTTGCTTTAGG  2040
Qy      2041  ATGACAAAATTTCAAGTACTCAGACTACAGATGTGTGACGGCCCTTCTGCTGCTG  2100
Db      2041  ATGACAAAATTTCAAGTACTCAGACTACAGATGTGTGACGGCCCTTCTGCTGCTG  2100
Qy      2101  ACCGAGCAAGCGCAGGCGCTTGGGGTGAAGAGCTGAAGCTTAAACGATATCATGCC  2160
Db      2101  ACCGAGCAAGCGCAGGCGCTTGGGGTGAAGAGCTGAAGCTTAAACGATATCATGCC  2160
Qy      2161  ACCCTTCTTCAATGATTTGACCTTGACATATGACATCGCTTGTGAGAGTGTGAG  2220
Db      2161  ACCCTTCTTCAATGATTTGACCTTGACATATGACATCGCTTGTGAGAGTGTGAG  2220
Qy      2221  CGGTGAGATGACAGCCTGTGTGCGCCCACTGCTGCTGATGCTAACCATGTCTTCC  2280
Db      2221  CGGTGAGATGACAGCCTGTGTGCGCCCACTGCTGCTGATGCTAACCATGTCTTCC  2280
Qy      2281  CTGCTGGCAAGGCAATCTGGGTGACAGGCTGGGGGCAACAAAGAGGATACCGAG  2340
Db      2281  CTGCTGGCAAGGCAATCTGGGTGACAGGCTGGGGGCAACAAAGAGGATACCGAG  2340
Qy      2341  CGCTGATCTGCAAGAGGTGAGATCCGTGATCAACAGACCACTGTGAGAGCTGA  2400
Db      2341  CGCTGATCTGCAAGAGGTGAGATCCGTGATCAACAGACCACTGTGAGAGCTGA  2400
Qy      2401  TGCCTGAGAGATCAACCCCAAGATATGTGTGCTGCTGCTGCTGCTGCTGCTG  2460
Db      2401  TGCCTGAGAGATCAACCCCAAGATATGTGTGCTGCTGCTGCTGCTGCTGCTG  2460
Qy      2461  CCTGCAAGGATGATCTGTGTGCGCTTGTCAAGCGCCGAGAAAGATGGCCGAT  2520
Db      2461  CCTGCAAGGATGATCTGTGTGCGCTTGTCAAGCGCCGAGAAAGATGGCCGAT  2520
Qy      2521  AGGCTGTGTGTGTGAGCTGGGGTGAAGCTGCTGCAAGAGAAAGAGGCGGTGTA  2580
Db      2521  AGGCTGTGTGTGTGAGCTGGGGTGAAGCTGCTGCAAGAGAAAGAGGCGGTGTA  2580
Qy      2581  CAAAGCTCCCTGTGTGTGCTGGGACCTGATCAAGAGACCACTGGGTATGACAG  2640
Db      2581  CAAAGCTCCCTGTGTGTGCTGGGACCTGATCAAGAGACCACTGGGTATGACAG  2640
Qy      2641  AGACAGCCGACCAACCAACCAAGAGATGCGGACCACTGATGATCAGAGAG  2700
Db      2641  AGACAGCCGACCAACCAACCAAGAGATGCGGACCACTGATGATCAGAGAG  2700
Qy      2701  GAAACATGACGATTTATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  2760
Db      2701  GAAACATGACGATTTATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  2760
Qy      2761  GCATCTTGAAGATCTGAGAGTCTTCAAAAGTGTGAGAGAGAGAGAGAG  2820
Db      2761  GCATCTTGAAGATCTGAGAGTCTTCAAAAGTGTGAGAGAGAGAGAGAGAG  2820
Qy      2821  TTGCGTGTGAGAGGCGGCGCTGGGGGCAAGGCTTGAAGAGAGAGAGAGAG  2880
Db      2821  TTGCGTGTGAGAGGCGGCGCTGGGGGCAAGGCTTGAAGAGAGAGAGAGAG  2880
Qy      2881  CTGAGCTGTGAGAGATGATGCTGCTGCGAGAGAGCTGCTCAACTGTCAATGAG  2940
Db      2881  CTGAGCTGTGAGAGATGATGCTGCTGCGAGAGAGCTGCTCAACTGTCAATGAG  2940
Qy      2941  GGGAGCCCTATGAGAGAGGCGCTCAGGCTCACTTTTCAAGAGAGAGAGAG  3000
Db      2941  GGGAGCCCTATGAGAGAGGCGCTCAGGCTCACTTTTCAAGAGAGAGAGAG  3000
Qy      3001  ACCCGAAGAAAGAGGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG  3060
Db      3001  ACCCGAAGAAAGAGGATGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG  3060
Qy      3061  TTGAGAGTAAACATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTT  3106
Db      3061  TTGAGAGTAAACATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTT  3106

```

## RESULT 2

AF042822

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## ORIGIN

AKKATWVMEKTHKGGTGAALLI0KGE1RVIN0TTCEDILPMQ0ITPRMVCVFLSGV  
 DSCGDSGSPUSAEKDRMP0AGVVSWE6GCAQKRNKPVYTRLRPVMDMKEHNGV"

Query Match 100.0%; Score 3106; DB 10; Length 3106;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGTGAAACGCGTCCCGAGAGGACCAAGCGCTGTAGACACCGGCATTCGGACCCGCCAAA 60  
 DB 1 CATGTGAAACGCGTCCCGAGAGGACCAAGCGCTGTAGACACCGGCATTCGGACCCGCCAAA 60  
 QY 61 CCATGGGTGAGCAATCGGGGCGCGAAGCGGAGGGGCTCTCAGACTTCGGGCGGAGAC 120  
 DB 61 CCATGGGTGAGCAATCGGGGCGCGAAGCGGAGGGGCTCTCAGACTTCGGGCGGAGAC 120  
 QY 121 TCAGTACAACTCCCGCTAGAGACATGATGCTTTGAGAGGGTGTGAGATTCTTCG 180  
 DB 121 TCAGTACAACTCCCGCTAGAGACATGATGCTTTGAGAGGGTGTGAGATTCTTCG 180  
 QY 181 CTGCGAACAATGCCAAGAAAGTGAAGAGCGAGCCCGACGCGCTGGGTGTGCTGTG 240  
 DB 181 CTGCGAACAATGCCAAGAAAGTGAAGAGCGAGCCCGACGCGCTGGGTGTGCTGTG 240  
 QY 241 CAGTGTCTTCACT 300  
 DB 241 CAGTGTCTTCACT 300  
 QY 301 ATTATCGAATGTGCGGGTCAAAAAGTCTTCAATGGCCATCTGAGAGTCACAATGAGA 360  
 DB 301 ATTATCGAATGTGCGGGTCAAAAAGTCTTCAATGGCCATCTGAGAGTCACAATGAGA 360  
 QY 361 TCTTTCTGATGGGTATGAGAACTCCACTCCACAGATTATTCAGCTTGGCCAGCCAG 420  
 DB 361 TCTTTCTGATGGGTATGAGAACTCCACTCCACAGATTATTCAGCTTGGCCAGCCAG 420  
 QY 421 TGAAGAGGGGCTGAGACCTGTGACATGAAAGTCCCTGCTGAGGCTCCCTACCAAGA 480  
 DB 421 TGAAGAGGGGCTGAGACCTGTGACATGAAAGTCCCTGCTGAGGCTCCCTACCAAGA 480  
 QY 481 AGTGGGCTGTAATCTGCTTCAAGTGAAGGCACTGATCATGCGCTACTACTGCTCAGAGTTCA 540  
 DB 481 AGTGGGCTGTAATCTGCTTCAAGTGAAGGCACTGATCATGCGCTACTACTGCTCAGAGTTCA 540  
 QY 541 GCATCCCCCACAACCTGCGACAGAGAGGTGATGCGCGCATGCTGAGAGGAGTTGTGA 600  
 DB 541 GCATCCCCCACAACCTGCGACAGAGAGGTGATGCGCGCATGCTGAGAGGAGTTGTGA 600  
 QY 601 CATTCGCAACCCCGAGACAGGGCACTGAAATCTTCTGATCTAACAATCTGTGGTGGCTTCC 660  
 DB 601 CATTCGCAACCCCGAGACAGGGCACTGAAATCTTCTGATCTAACAATCTGTGGTGGCTTCC 660  
 QY 661 CCATTGACCCCGAGATGTGACAGAGACTCAGACACACAGCTGCAGTTTGCCTGCATG 720  
 DB 661 CCATTGACCCCGAGATGTGACAGAGACTCAGACACACAGCTGCAGTTTGCCTGCATG 720  
 QY 721 CCCATGTGTGACAGTGAACGCTTCACTAACCCCTGGCTTCCCAACAGTCCCTACCCGG 780  
 DB 721 CCCATGTGTGACAGTGAACGCTTCACTAACCCCTGGCTTCCCAACAGTCCCTACCCGG 780  
 QY 781 CGCATGCCCCGCTGCGAGTGGTCTGCGGGGGGAGCGCGACTCTGTGTGAGGCTCACT 840  
 DB 781 CGCATGCCCCGCTGCGAGTGGTCTGCGGGGGGAGCGCGACTCTGTGTGAGGCTCACT 840  
 QY 841 TCCGAGCTTTGATGTGCTTCCCTGTATGAGCATGGAAGTGAAGCTGTGACCTGTATG 900  
 DB 841 TCCGAGCTTTGATGTGCTTCCCTGTATGAGCATGGAAGTGAAGCTGTGACCTGTATG 900  
 QY 901 ATAGCTTGAAGCCCATGGAACCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
 DB 901 ATAGCTTGAAGCCCATGGAACCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
 QY 961 CCTACAACTGACTTCTCTCTCTCCAGAAAGTCTTCTCTTGTCAAGCTGATTAACAAATA 1020

DB 961 CCTACAACTGACTTCTCTCTCTCCAGAAAGTCTTCTCTTGTCAAGCTGATTAACAAATA 1020  
 QY 1021 CTGACCGGAGCATCTGCTTGAAGGCACTTCTTCAGCTGCGCCAGATGAGCACT 1080  
 DB 1021 CTGACCGGAGCATCTGCTTGAAGGCACTTCTTCAGCTGCGCCAGATGAGCACT 1080  
 QY 1081 GTGGGCGCTTTTGAAGTGAACCCCAAGGACATTTAGACGCGCTTACTATCCAGGCACT 1140  
 DB 1081 GTGGGCGCTTTTGAAGTGAACCCCAAGGACATTTAGACGCGCTTACTATCCAGGCACT 1140  
 QY 1141 ACCCGCCAACTCACTGACATGGAATATCAAGGTGCCCAACACCGGAAAGTGAAG 1200  
 DB 1141 ACCCGCCAACTCACTGACATGGAATATCAAGGTGCCCAACACCGGAAAGTGAAG 1200  
 QY 1201 TGGGCTTCAAACTCTTCTATCTGTGGAACCCCAAGTACAGTGGGCTCTCTGACCAAG 1260  
 DB 1201 TGGGCTTCAAACTCTTCTATCTGTGGAACCCCAAGTACAGTGGGCTCTCTGACCAAG 1260  
 QY 1261 ACTATGTGAGATCAACGGGAGAACTGCGGTGAGAGAGTCCAGTTGTGTGAGCA 1320  
 DB 1261 ACTATGTGAGATCAACGGGAGAACTGCGGTGAGAGAGTCCAGTTGTGTGAGCA 1320  
 QY 1321 GCAACAGCAGCAAGATTAAGTCCACTTCCATCTGATCACTGTACAGGACACCGGGT 1380  
 DB 1321 GCAACAGCAGCAAGATTAAGTCCACTTCCATCTGATCACTGTACAGGACACCGGGT 1380  
 QY 1381 TCTTAGGTGATGACTCTCTCTTCAAGACCTCAACAGCCGTCGAGAGATGTTATGATGCA 1440  
 DB 1381 TCTTAGGTGATGACTCTCTCTTCAAGACCTCAACAGCCGTCGAGAGATGTTATGATGCA 1440  
 QY 1441 AGACTGAGCGGTGATCCGAAAGAACTGCGGTGAGAGAGTCCAGTTCAGTGCAG 1500  
 DB 1441 AGACTGAGCGGTGATCCGAAAGAACTGCGGTGAGAGAGTCCAGTTCAGTGCAG 1500  
 QY 1501 ATAGTATGAGCGGTATCTGCGCATGGAATGCCACCCAGCAAGTTCAGTGCAG 1560  
 DB 1501 ATAGTATGAGCGGTATCTGCGCATGGAATGCCACCCAGCAAGTTCAGTGCAG 1560  
 QY 1561 TCTGCAAGCCCTCTTCTGAGGTCTGTGACAGTGTCAACGACTGTGGGAGCGAAAGTGA 1620  
 DB 1561 TCTGCAAGCCCTCTTCTGAGGTCTGTGACAGTGTCAACGACTGTGGGAGCGAAAGTGA 1620  
 QY 1621 AGAGGCGCTGAGCTGCTCTGCTGAGGATTTCAAGTGTTCATATGGAAGTGTCCCTC 1680  
 DB 1621 AGAGGCGCTGAGCTGCTCTGCTGAGGATTTCAAGTGTTCATATGGAAGTGTCCCTC 1680  
 QY 1681 AGAGCCGAAGTGTAAATGGGAAGGACAACTGTGAGATGGGTCTGACGAGGCTTCATGTG 1740  
 DB 1681 AGAGCCGAAGTGTAAATGGGAAGGACAACTGTGAGATGGGTCTGACGAGGCTTCATGTG 1740  
 QY 1741 ACAGCGTAATGTGCTCTCTTGCACCAATATTAACCTACCGCTGCAAAATGGCTCTGTG 1800  
 DB 1741 ACAGCGTAATGTGCTCTCTTGCACCAATATTAACCTACCGCTGCAAAATGGCTCTGTG 1800  
 QY 1801 TGAGCAAGGCAACCTGAGTGTATGAGGAAGAGAGACTGTGAAGATGGCTCCGATGAGA 1860  
 DB 1801 TGAGCAAGGCAACCTGAGTGTATGAGGAAGAGAGACTGTGAAGATGGCTCCGATGAGA 1860  
 QY 1861 AAAACTGTGATGTGGGCTGCGATCTTTACCAACAGAGCTCGAGTGTGTGTGTGTGTGT 1920  
 DB 1861 AAAACTGTGATGTGGGCTGCGATCTTTACCAACAGAGCTCGAGTGTGTGTGTGTGTGT 1920  
 QY 1921 ATGCGAAGAGGCGAGAGTGGGCTGTGAGAGCTTCACGCTTCGAGGCAAGGCGCACT 1980  
 DB 1921 ATGCGAAGAGGCGAGAGTGGGCTGTGAGAGCTTCACGCTTCGAGGCAAGGCGCACT 1980  
 QY 1981 TGTGTGGGCGCTCTGCTCATCTCTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040  
 DB 1981 TGTGTGGGCGCTCTGCTCATCTCTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040  
 QY 2041 ATGCAAAATTTCAAGTACTCAGACTACAGATGGAAGGCTTCTGAGTGTGTGTGTGTGT 2100







QY 1321 GCAACAGCAGCAAGATTACAGTCCATTCCATTCTGATCACTGCTACACGACACCCGGT 1380  
 DB 1431 GCAACAGCAGCAAGATTACAGTCCATTCCATTCTGATCACTGCTACACGACACCCGGT 1490  
 QY 1381 TCTAGCTGATGATCTCTCTCTCAAGCTCAACGACCCGCTGCCAGGAGATTGATGATGCA 1440  
 DB 1491 TCTAGCTGATGATCTCTCTCTCAAGCTCAACGACCCGCTGCCAGGAGATTGATGATGCA 1550  
 QY 1441 AGACTGAGCGGTGATTCGCAAGAAAGAACTGCTGCTGCAAGCGCTGCGGAGACTGCCCCGAT 1500  
 DB 1551 AGACTGAGCGGTGATTCGCAAGAAAGAACTGCTGCTGCAAGCGCTGCGGAGACTGCCCCGAT 1610  
 QY 1501 ATAGTATGAGCGTTACTGCGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1560  
 DB 1611 ATAGTATGAGCGTTACTGCGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1670  
 QY 1561 TCTGCAAGCCCTCTCTCTGAGTCTGATGACAGTCTCAACGACTGCTGCGGAGACGAAAGTGA 1620  
 DB 1671 TCTGCAAGCCCTCTCTCTGAGTCTGATGACAGTCTCAACGACTGCTGCGGAGACGAAAGTGA 1730  
 QY 1621 AGGAGGCTGACAGTCTCTCTCTGAGTCTGATGACAGTCTCAACGACTGCTGCGGAGACGAA 1680  
 DB 1731 AGGAGGCTGACAGTCTCTCTCTGAGTCTGATGACAGTCTCAACGACTGCTGCGGAGACGAA 1790  
 QY 1681 AGAGCCAGAGTGTATGAGAAAGAAAGCAACGCTGAGATGAGTGTGACGAGCTTCATGTC 1740  
 DB 1791 AGAGCCAGAGTGTATGAGAAAGAAAGCAACGCTGAGATGAGTGTGACGAGCTTCATGTC 1850  
 QY 1741 ACAGCGTGAATGTCTCTCTCTGACCCAAATATATCTACCGCTGCGCAAAATGAGCTCTGTC 1800  
 DB 1851 ACAGCGTGAATGTCTCTCTCTGACCCAAATATATCTACCGCTGCGCAAAATGAGCTCTGTC 1910  
 QY 1801 TGAAGCAAGGCGCAACCTGATGATGAGAAAGCAAGCTGAGCGATGAGCTCCGATGAGA 1860  
 DB 1911 TGAAGCAAGGCGCAACCTGATGATGAGAAAGCAAGCTGAGCGATGAGCTCCGATGAGA 1970  
 QY 1861 AAAACGTGATCTGTGGGCTGCGGATCTCTTACCAAAACAGCTGCGGTGTGTGGACGA 1920  
 DB 1971 AAAACGTGATCTGTGGGCTGCGGATCTCTTACCAAAACAGCTGCGGTGTGTGGACGA 2030  
 QY 1921 ATGCGGACGAGGCGAGTGGCCCTGCGAGTGAAGCTCCACGCGCTGCGGCGAGGCGCACT 1980  
 DB 2031 ATGCGGACGAGGCGAGTGGCCCTGCGAGTGAAGCTCCACGCGCTGCGGCGAGGCGCACT 2090  
 QY 1981 TGTGTGGGCGCTGCTCATCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
 DB 2091 TGTGTGGGCGCTGCTCATCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2150  
 QY 2041 ATGACAAAATTTCAAGTACTGAGACTACAGATGATGAGTGGAGCGGCTTCTGCGGTCTGCTGG 2100  
 DB 2151 ATGACAAAATTTCAAGTACTGAGACTACAGATGATGAGTGGAGCGGCTTCTGCGGTCTGCTGG 2210  
 QY 2101 ACCAGAGCAAGCGAGTCTCTGCGGCTGCGAGAGCTGAAAGCTCAAAAGTATCATCAACC 2160  
 DB 2211 ACCAGAGCAAGCGAGTCTCTGCGGCTGCGAGAGCTGAAAGCTCAAAAGTATCATCAACC 2270  
 QY 2161 ACCCTTCTTCAATGATTTACCTTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
 DB 2271 ACCCTTCTTCAATGATTTACCTTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2330  
 QY 2221 CGGTGAGATCAGACCGTGTGGGCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
 DB 2331 CGGTGAGATCAGACCGTGTGGGCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2390  
 QY 2281 CTGCTGAGAGCGCATCTGAGTCAAGGCTGCGGCGACACAAAAGAGGAGGTACCGGAG 2340  
 DB 2391 CTGCTGAGAGCGCATCTGAGTCAAGGCTGCGGCGACACAAAAGAGGAGGTACCGGAG 2450  
 QY 2341 CGCTGATCTTGCAGAAAGGCTGATCGTGTCTATCAACGAGACCACTGTGAGGACCTCA 2400  
 DB 2451 CGCTGATCTTGCAGAAAGGCTGATCGTGTCTATCAACGAGACCACTGTGAGGACCTCA 2510  
 QY 2401 TGCCGAGCAGATATCAACCCAGATGATGTGTGTGGGTTCTCTGAGTGGGCTGTGAGCT 2460

DB 2511 TGCCGAGCAGATATCAACCCAGATGATGTGTGGTTTCTCAAGTGGGCTGTGAGCT 2570  
 QY 2461 CTGCGCAGGCTGATCTGTGTGGCTCTTGTCAAGCCGCGAGAAAGATGCGATGTTCC 2520  
 DB 2571 CTGCGCAGGCTGATCTGTGTGGCTCTTGTCAAGCCGCGAGAAAGATGCGATGTTCC 2630  
 QY 2521 AGGCTGATGTGTGAGCTGTGGGCTGAAAGGCTGCGCTGAGAGGAAACAGAGGCTGTGACA 2580  
 DB 2631 AGGCTGATGTGTGAGCTGTGGGCTGAAAGGCTGCGCTGAGAGGAAACAGAGGCTGTGACA 2690  
 QY 2581 CAAAGCTCCTGTAGTTGCGGACTGATCAAAAGAGCACTGCGGTATATACAGATGAGAC 2640  
 DB 2691 CAAAGCTCCTGTAGTTGCGGACTGATCAAAAGAGCACTGCGGTATATACAGATGAGAC 2750  
 QY 2641 AGACAGCCAGCCACCAAAACCCACAGAGATGCCGATGACACACTGATACAGAGAG 2700  
 DB 2751 AGACAGCCAGCCACCAAAACCCACAGAGATGCCGATGACACACTGATACAGAGAG 2810  
 QY 2701 GAACACTGACGACATTTATGCTGTGGCTGCCCGCCCGCAACCAACCCAGACTGTGACT 2760  
 DB 2811 GAACACTGACGACATTTATGCTGTGGCTGCCCGCCCGCAACCAACCCAGACTGTGACT 2870  
 QY 2761 GCATCCTTGAAGCTCAGAGTCTTCCAAAGTGGAGCCCTCAAGAGTTGGAGAGAGAAC 2820  
 DB 2871 GCATCCTTGAAGCTCAGAGTCTTCCAAAGTGGAGCCCTCAAGAGTTGGAGAGAGAAC 2930  
 QY 2821 TTGCGGTCTAGCGGCGCCAGCTGCGGCGCAAGGTTTATGAGGACCTTCCCTTACGCC 2880  
 DB 2931 TTGCGGTCTAGCGGCGCCAGCTGCGGCGCAAGGTTTATGAGGACCTTCCCTTACGCC 2990  
 QY 2881 CTGAGCTGGGTGAGATGATGTGTGCTGCCGAGAGCTGCTTCCACTGTATGAGTCCC 2940  
 DB 2991 CTGAGCTGGGTGAGATGATGTGTGCTGCCGAGAGCTGCTTCCACTGTATGAGTCCC 3050  
 QY 2941 GCGAGCCCTTATGAGAGAGGCGCTCAGGCTCACTCTTTTCAGAGAACCCGACCTTACAGA 3000  
 DB 3051 GCGAGCCCTTATGAGAGAGGCGCTCAGGCTCACTCTTTTCAGAGAACCCGACCTTACAGA 3110  
 QY 3001 ACCCGAGAAAGAGTGTATACCTTAAGCTGGAATTTGTTTCTGCTTCCAGGCGGTGAT 3060  
 DB 3111 ACCCGAGAAAGAGTGTATACCTTAAGCTGGAATTTGTTTCTGCTTCCAGGCGGTGAT 3170  
 QY 3061 TTGAGAGTAAACATTTATTTCTTTTAAAAAAA 3106  
 DB 3171 TTGAGAGTAAACATTTATTTCTTTTAAAAAAA 3216  
 RESULT 4  
 AB049189 3174 bp mRNA linear ROD 14-APR-2001  
 LOCUS  
 DEFINITION Rattus norvegicus mRNA for membrane bound arginine specific serine  
 protease, complete cds.  
 AB049189  
 ACCESSION  
 VERSION AB049189.1 GI:10336526  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE  
 1 (sites)  
 Inoue, H., Takahashi, K. and Kishi, K.  
 membrane-bound arginine specific serine protease  
 Published Only in Database (2000)  
 2 (bases 1 to 3174)  
 Inoue, H., Takahashi, K. and Kishi, K.  
 Direct Submission  
 Submitted (22-SEP-2000) Hideaki Inoue, Tokyo University of Pharmacy  
 and Life Science, School of Life Science, 1432-1 Horiinouchi,  
 Hachioji-shi, Tokyo 192-0392, Japan (E-mail: hinoue@is.toyaku.ac.jp,  
 Tel: 81-426-76-7153, Fax: 81-426-76-7157)  
 Location/Qualifiers  
 FEATURES

```

source
1. .3174
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="duodenum"
/dev_stage="7 weeks old"
127..2694
/codon_start=1
/evidence=experimental
/product="membrane bound arginine specific serine
protease"
/protein_id="BAB13765.1"
/db_xref="GI:10336527"
/translation="MGNNRGRKAGGGSDPFGALIKYNSLNNNGEBSVEFLPVNNA
KQVEKGRPRMVAIVAFSLIMAGLWHPHYRNRLOKXNGHILTNELF
DAVENTSTEFISLQVKEALMYSEVPLGYPHKSTVLAIFSGSVIAYWSFIS
IPPHLEERYDRAVVRVVTLPARALSPVLTSVAPIDRMLORTDONSFPAL
HARGTRVFTTPGPNSPYPAHRCQWTLRGDSDVLSLTPSPVACDGDSDLV
TVYDSLSEPHAVVRLCGTFSTNLTLSQNVFLVLTITDRHRHGFATPDL
PVMSGCGALISEQTFSSPYPGHYPNINCTWNLKPNRNVKRFKFLVLPDNI
PVGCTKDYVLEINKEKFCGRSQFVSSNSKLTVAHSDHSTDTGFLAEYISYDN
DPCGEMCKTGRCIRKLDQMDPCPYSDERHCRNATQPMCKNOFCRPLFWVC
DSVNDGSDGSDERGCSPAGSFCKSKCLPQOCCNGKDCGSDSDASCDVNAVS
CTKYTRCONGLINKNPPCDGKDCSDGSDKNCDCGLRSTKQARVVGTMADG
EWPQVSLHAGGHLGASLISPPMLVSAHCFODETYFKYSDHTMTAFLGLDLS
KRSASGVQEHKRLRIITHPSENDFTDYDIALLELKEPSEYSTVTRPCLPDTHYF
AGKAIWVTGHTKEGTGALILQKGBIRINQTEELLPQDITRRMNCVGLSGV
DSQGDSDGSPLSVEKGRIFQAGVSVWEGCAQRNKPQVYTRIPEVRMIKEQTV"
360
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
513
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
573
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
2121
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
3157..3174
polysite
ORIGIN
Query Match 83.5%; Score 2595; DB 10; Length 3174;
Best Local Similarity 91.1%; Pred. No. 0; Mismatches 230; Indels 50; Gaps 7;
Matches 2853; Conservative 0;

```

```

QY 301 ATTATGGAAATGTCGGGTTCAAAAAAGTCTTCATGGCCATCTGAGGATCACAAAATGAGA 360
DB 365 ATTACGGGAATGTAGGATCCAAAAAGTCTTCATGGCCATCTGAGGATCAAGAAATGAGA 424
QY 361 TCTTTCTGGATCGGATAGAACTCACCCTCCACAGATTTCAGCCCTGGCCAGCAGG 420
DB 425 ACTTTCTGGACCGATAGAACTCACCCTCCACAGATTTCAGCCCTGGCCAGCAGG 484
QY 421 TGAAGAGGCGTGAAGCTGTCAATGAAGTCCCTGCTGGTCCCTACCAAGA 480
DB 485 TGAAGAGGCGTGAAGCTGTCAATGAAGTCCCTGCTGGTCCCTACCAAGA 544
QY 481 AGTCGCTGAACTGCTTCAGTGAAGGCAGTGTATCGCTTACTACTGTGACAGATTCA 540
DB 545 AATGACTTAACTGCTTCAGTGAAGGCAGTGTATCGCTTACTACTGTGACAGATTCA 604
QY 541 GCATCCCCCAACCTGGAGAAAGTTGATGGGCCCATGGCTGGAGGCGAGTTGTA 600
DB 605 GCATCCCCCAACCTGGAGAAAGTTGATGGGCCCATGGCTGGAGGCGAGTTGTA 664
QY 601 CATTGCAACCCGAGACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGGCTTCC 660
DB 665 CACTGCAACCCGAGACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGGCTTCC 724
QY 661 CCATTGACCCAGATGCTGACAGAGACTCAGAGACAACAGTGCAGTTTGGCTGAGT 720
DB 725 CCATTGACCCAGATGCTGACAGAGACTCAGAGACAACAGTGCAGTTTGGCTGAGT 784
QY 721 CCCATGCTGAGAGAGACAGCTTCACTACCTCCCTGCTCCCAAGTCCCTACCCG 780
DB 785 CCCGTCAGAGAGAGACAGCTTCACTACCTCCCTGCTCCCAAGTCCCTACCCG 844
QY 781 CGCATGCCGCTGCAAGTGGTCTGCGGGGAGACGCCGACTGTGCTGAGCTTCACT 840
DB 845 CTGATGCCGCTGCAAGTGGTCTGCGGGGAGACGCCGACTGTGCTGAGCTTCACT 904
QY 841 TCCGAAGCTTGAATGTCGCTCCTGCTGATGACATGACATGACCTGATACCGTATG 900
DB 905 TCCGAAGCTTGAATGTCGCTCCTGCTGATGACATGACATGACCTGATACCGTATG 964
QY 901 ATGAGCTGAGCCCATGAGAACCCCAAGCTGAGGCGCTGTGAGCACTTTCACCT 960
DB 965 ACAGCTGAGCCCATGAGAACCCCAAGCTGAGGCGCTGTGAGCACTTTCACCT 1024
QY 961 CTTACAACTGACTTCTCTCTCCCAAGAGCTTCTTTGTCAAGCTGATTAACAATA 1020
DB 1025 CTTACAACTGACTTCTCTCTCCCAAGAGCTTCTTTGTCAAGCTGATTAACAATA 1084
QY 1021 CTGACCGGCGCATCTGCTGCTTGAAGCCACTTCTTCCAGCTGCCAAGATGAGCAGCT 1080
DB 1085 CTGACCGGCGCATCTGCTGCTTGAAGCCACTTCTTCCAGCTGCCAAGATGAGCAGCT 1144
QY 1081 GTGGCGGCTTTTGAAGGACCAACCAAGGACATTTACAGGCCCTACTATCCAGGCCACT 1140
DB 1145 GTGGCGGCTTTTGAAGGACCAACCAAGGACATTTACAGGCCCTACTATCCAGGCCACT 1204
QY 1141 ACCGCGCCCAATCACTGACATGAAATATCAAGGTGCCCCAACAACCGGAAGCTGAAG 1200
DB 1205 ACCGCGCCCAATCACTGACATGAAATATCAAGGTGCCCCAACAACCGGAAGCTGAAG 1264
QY 1201 TGGCTTCAAACTCTTCTATCTGCTGACCCCAACGATACGATGAGGCTCTGCAACAAG 1260
DB 1265 TACGCTTCAAACTCTTCTATCTGCTGACCCCAACGATACGATGAGGCTCTGCAACAAG 1324
QY 1261 ACTATGAGGATCAACGGGAGAAAGTACTCGGTGAGAGGCTCCAGTTTGTGTAGCA 1320
DB 1325 ACTATGAGGATCAACGGGAGAAAGTACTCGGTGAGAGGCTCCAGTTTGTGTAGCA 1384
QY 1321 GCACAGCAGCAAGATTACAGTCACTTTCATCTGATCACTGTAACGACACCGGGT 1380
DB 1385 GCACAGCAGCAAGATTACAGTCACTTTCATCTGATCACTGTAACGACACCGGGT 1444
QY 1381 TCTTAGCTGAGTACCTCTCTTAACGACTCAACGACCGGTGCCAAGGATGTTCAATGTGA 1440

```



```

Db      1445  TCCGTGGTGAATCTCTCTTACGACTCCATGACCCCATGGATGGATGATGTA 1504
Qy      1441  AGACTGGAAGGTCATCCGAAAGAACTGGCGTCGCGAGAGCTGGGAGACTGGCCGATT 1500
Db      1505  AGACTGGAAGGTCATCCGAAAGAACTGGCGTCGCGAGAGCTGGGAGACTGGCCGATT 1564
Qy      1501  ATGTGATGAGCGTTACTGCGCATGCAATCCGACCCGACGTTCACTGCAAAAAACAGT 1560
Db      1565  ACAGTATGAGCGTTACTGCGCATGCAATCCGACCCGACGTTCACTGCAAAAAACAGT 1624
Qy      1561  TCTGCAAGCCCTCTTCTGGGTCGTGACAGTGTCAACGCTGTGGGAGCGAAAGTACG 1620
Db      1625  TCTGCAAGCCCTCTTCTGGGTCGTGACAGTGTCAACGCTGTGGGAGCGAGGATG 1684
Qy      1621  AGGAGGCTGAGCTGTCTGCTGGGAGTTTCAAGTGTCCATGGGAAAGTGTCCCTC 1680
Db      1685  AGGAGGCTGAGCTGTCTGCTGGGAGTTTCAAGTGTCCATGGGAAAGTGTCCCTC 1744
Qy      1681  AGAGCGAAGATCTATGGAAGAGCACTGTGAGATGGGTGTGACGAGGCTTCATGTG 1740
Db      1745  AGAGCGAAGATCTATGGAAGAGCACTGTGAGATGGGTGTGACGAGGCTTCATGTG 1804
Qy      1741  ACAGCGTGAATGCTGCTCTTGGACCAAAATATACCTGACGCTGGCCAAATGGCTGTG 1800
Db      1805  ACAGTGTGATGCGCTCTTGTGACCAAAATATACCTGACGCTGGCCAAATGGCTGTG 1864
Qy      1801  TGAAGCAAGGCAACCTGATGTGATGGAAGACGAGCTGTGAGCATGAGTCCGATAGA 1860
Db      1865  TGAAGCAAGGCAACCTGATGTGATGGAAGACGAGCTGTGAGCATGAGTCCGATAGA 1924
Qy      1861  AAAACGTGACCTGTGGGCTGCGATCTTTTACCAAAAGGCTGTGCTGTGGTGGACGA 1920
Db      1925  AAAACGTGACCTGTGGGCTGCGATCTTTTACCAAAAGGCTGTGCTGTGGTGGACGA 1984
Qy      1921  ATGCGGACGAGGGGAGTGGCCCTGGGAGGTGAGGCTCCAGCCCTGGGCGGAGGCTACT 1980
Db      1985  ATGCGGACGAGGGGAGTGGCCCTGGGAGGTGAGGCTCCAGCCCTGGGCGGAGGCTACT 2044
Qy      1981  TGTGTGGGCTCTGCTCATCTCTCTGATGATGCTGGTCTGTGACAGTCAATGCTTTACG 2040
Db      2045  TGTGTGGGCTCTGCTCATCTCTCTGATGATGCTGGTCTGTGACAGTCAATGCTTTACG 2104
Qy      2041  ATGACAAAAATTTCAAGTACTCAAGACTACAGATGTGAGCGGCTTCTGGGCTGTGCTG 2100
Db      2105  ATGACAAAAATTTCAAGTACTCAAGACTACAGATGTGAGCGGCTTCTGGGCTGTGCTG 2164
Qy      2101  ACCAGAGCAAGCGAGTGCCTCTGGGCTGCAAGCTGAAGCTCAAGTATCATCAACC 2160
Db      2165  ACCAGAGCAAGCGAGTGCCTCTGGGCTGCAAGCTGAAGCTCAAGTATCATCAACC 2224
Qy      2161  ACCCTTCTTCAATGATTTCACTTGAATGATGATGCGCTTGTGAGCTGAGAGAT 2220
Db      2225  ACCCTTCTTCAATGATTTCACTTGAATGATGATGCGCTTGTGAGCTGAGAGAT 2284
Qy      2221  CGGTGAGTACAGACCGTCTGGGCGCCCACTGCTGCTGCTGATGCTAACATGCTTTC 2280
Db      2285  CGGTGAGTACAGACCGTCTGGGCGCCCACTGCTGCTGCTGATGCTAACATGCTTTC 2344
Qy      2281  CTGCTGGAAGGATCTGGGTCAAGGCTGGGGGACACAAAAGAGGAGGATACCGGAG 2340
Db      2345  CTGCTGGAAGGATCTGGGTCAAGGCTGGGGGACACAAAAGAGGAGGATACCGGAG 2404
Qy      2341  CGCTGATCTTGCAGAAAGGATGATCCGTGTATCAACACAGACCACTGTGAGACCTCA 2400
Db      2405  CACTGATCTTGCAGAAAGGATGATCCGTGTATCAACACAGACCACTGTGAGACCTCT 2464
Qy      2401  TGCGGAGGAGATACCCCAAGATGATGTGTGGGTTTCTCACTGGGGGCTGTGACT 2460
Db      2465  TGCGGAGGAGATACCCCAAGATGATGTGTGGGTTTCTCACTGGGGGCTGTGACT 2524
Qy      2461  CTTGCAAGGATCTGTGGGCTTGTGCAAGCGGAGAAAGATGGGGAATGTTCC 2520

```

```

Db      2525  CTTGCCAGGGTGAATCCGGTGCCCTTGTTCAGTGTGAGAAAGATGGGCAATCTTCC 2584
Qy      2521  AGGCTGTGTGTGAGCTGGGGTGAAGGCTTGGCTCAGAGAAACAGCGCTGTACA 2580
Db      2585  AGGCTGTGTGTGAGCTGGGGTGAAGGCTTGGCTCAGAGAAACAGCGCTGTACA 2644
Qy      2581  CAAGGCTCCCTGTGTGTGAGCTGGGATCAAGAGCACTGGGGTGTATACAGATGAC 2640
Db      2645  CAAGGATCCCTGAAGTGTGGGACTGAGTCAAAAGCAAACTGGGATGTAAAGATGAC 2704
Qy      2641  AGACAGCCGACCAAAACACCAAGAGATGCCGACATGACACACTGATACAGAGAG 2700
Db      2705  AGACAAACCAACCATGAATCTGCAAGAGATGCCGCCGACGATCTGATACAGAGAG 2764
Qy      2701  GAACACTGACGATTTATGTGTGGCTTCCCCCCCCAACACACACCCAGACTGTGAAT 2760
Db      2765  GATGCTGAGGATATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2819
Qy      2761  GCATCTTGAAGTCTGAGTCTTCA-----AGTGGGACCCCTCAGAGTTGG 2811
Db      2820  GCTTCCCTGAGTCTGAGATCTTCCCTGAGCTTCCAAAGTGGAGCCCTCAGAGTTGG 2879
Qy      2812  AAGAGAACTTGTGCTGACGAGCCGACGCTG---GGGCAAGGTTGATGAGCAGCT 2867
Db      2880  CCAAGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2939
Qy      2868  TCCCTCTGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2922
Db      2940  T-CCCTCTGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2998
Qy      2923  -----AAGTCAATGAGCTCCCGGAGCCCTATGAGGAGGAGGAGGCTCAGGCTAC 2973
Db      2999  CTAAGAAAGAACTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3054
Qy      2974  TCTTTTCAAGAAAGCCGACGCTTGAAGAACCCCAAGAAAGAGTGTACCTTAAGCTGAAT 3033
Db      3055  GTCTTTCAAGAAAGCTTCAAGCCCTTGAAGAACCCCAAGAAAGAGTGTACCTTAAGT 3114
Qy      3034  TGTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3093
Db      3115  TGTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3161
Qy      3094  AAAAAAAAAAAAAA 3106
Db      3162  AAAAAAAAAAAAAA 3174

```

```

RESULT 5
AB037898
LOCUS   2568 bp mRNA linear ROD 20-OCT-2001
DEFINITION
AB037898 Rattus norvegicus MBSP mRNA for membrane bound serine protease,
complete cds.
ACCESSION
AB037898
VERSION
AB037898.1 GI:9650963
KEYWORDS
Membrane bound serine protease; MT-SPL.
SOURCE
Rattus norvegicus (Norway rat).
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
Satomi,S., Yamaaki,Y., Tsuzuki,S., Hiromi,Y., Imanaga,T. and
Fushiki,T.
A role for membrane-type serine protease (MT-SPL) in intestinal
epithelial turnover
Biochem. Biophys. Res. Commun. 287 (4), 995-1002 (2001)
MEDLINE
21458307
PUBMED
11573963
2 (bases 1 to 2568)
REFERENCE
AUTHORS
Tsuzuki,S.
TITLE
Direct Submission
SUBMITTED (26-JUN-2000) Satoshi Tsuzuki, Kyoto University, Graduate
School of Agriculture, Division of Food Science and Biotechnology.
JOURNAL

```

Laboratory of Nutrition Chemistry, Oiwake-cho, Kitashirakawa,  
Sakyo-ku, Kyoto 606-8502, Japan  
(E-mail: tkhono@kais.kais.kyoto-u.ac.jp, Tel: 81-75-753-6263,  
Fax: 81-75-753-6264)

FEATURES  
source Location/Qualifiers

1. 2568  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/sex="male"  
/tissue\_type="jejunum"  
/dev\_stage="8 weeks old"  
1. 2568  
/gene="MBSP"  
1. 2568  
/gene="MBSP"  
/codon\_start=1  
/product="membrane bound serine protease"  
/protein\_id="BAB03502.1"  
/db\_xref="GI:9650964"  
/translation="MGNNRGRKAGGGSDPAGLKYNSRLNNMGFEVEEFLPVNNA  
/KVEKGRPRMVMVAIVFSEFLISLIMAGLVMHFRVNRVLRQKVENGLRTNENFL  
DAENSTSEPIELASOVKALMTATSEVLPYHKSIVTASEGSLIAYVMSDF  
IPHLEEDRAVAVERVVTIPRARALSFVLSVAVFIDPRMDORTDNCSFAL  
HARGTVTRPTTTPGPNSEVPARARCOMVIRGADSVLSLTERSPVAPDGDSDLV  
TVYDLSPEMPEHVAVRVLCGFSPSYNLTPISSQNVFLVLTITNDRHREFEATPOL  
PKMSGCGGLSEAOGTSSPYPGHYPNINCTWNIKVENNRVAKRFLFYIVDNT  
PVSCCTDYVEIENGKFCGERSQVVSNSKLTIFHSHDSHTDGLFVLAIVSYDN  
DPCKMCKTGKICRKLDCWADCPVDSRHRCAVTHQPMCKNFCCKLFWC  
CTSYNDGDSDESGCSGPAISFKNSGKCLPOSQCNKGDGSDGDEASCDNVAVS  
CTRYTRCCNGCLTANKNPECDKDCSDSDERKNDGCRSTKARVAVGNNADG  
EMWQVSLHNLGGHLCGASLISPDMLVSAHCFQDRTIKYSDHMTWTFGLILLOS  
KRASGVQEHKRLRIITHPSENFDTFYDIALLEKPAVSTVAPRICPDNTHRP  
AGAAIVTGWHTREGTGALILQKGRIVINQITCEELLPOQITPRMCKVGLSGV  
DSCQDSGGLSVEKGRIFQAGVWSMBCAQRNKPVYTRIPEVRDMIKQTV"

## ORIGIN

Query Match 74.5%; Score 2313.6; DB 10; Length 2568;

Best Local Similarity 93.8%; Pred. No. 0; Mismatches 159; Indels 0; Gaps 0;

Matches 2409; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

63 ATGGGTGCAATCGGGCGCGCAAGCGGAGGGGCTCTCAGACCTCGGCGCGGAGCTC 122  
Db 1 ATGGGAAACAATCGGGCGCGCAAGCGGAGGGGCTCGAGACCTCGGCGCGGAGCTC 60  
Qy 123 AAGTACAATCCCGGTAGAGACATGATGCTTTAGAGAGGCTGTGAGATTCCTGCT 182  
Db 61 AAGTACAATCCCGGTAGAGACATGATGCTTTAGAGAGGCTGTGAGATTCCTGCT 120  
Qy 183 GCGAACAATGCCAAGAGTGGAGAGGAGCCCGAGCGCTGGGTGGTGGTGGCA 242  
Db 121 GTGAACAACGCAAGCGGTGAGAGAGGAGCGCCCGAGCGCTGGGTGGTGGTGGC 180  
Qy 243 GTGCTGTAGCT 302  
Db 181 GTAGTGTCACT 240  
Qy 303 TATCGAATGTGCGGGTCAAAAAGTCTTCAATGGCCATCTGAGATCAAAAGATC 362  
Db 241 TACGGAAATGTGAGGATCCAAAAGTCTTCAATGGCCATCTGAGATCAAAAGATC 300  
Qy 363 TTTCTGATGGGTATGAGAACTCCACCTCCACAGGTTTATCAGCTGGCCAGCGAGTG 422  
Db 301 TTTCTGAGCGGTATGAGAACTCCACCTCCACAGGTTTATCAGCTGGCCAGCGAGTG 360  
Qy 423 AAGAGAGCGGTAGAGCTGTGACATGAGCTCTGTCTGGGCTCCCTACCAAGAAG 482  
Db 361 AAGAGAGCGGTAGAGCTGTGACATGAGCTCTGTCTGGGCTCCCTACCAAGAAG 420  
Qy 483 TCGGCTGTAGCTCTGTAGTGGAGGCACTGTGATCGCTTACTGTGTGAGATTCAGC 542  
Db 421 TCGAATGTAGCTCTGTAGTGGAGGCACTGTGATCGCTTACTGTGTGAGATTCAGC 480

Qy 543 ATCCCCCACCACCTGAGCAAGAGGTTGATCGGCCATGCTGTGAGCGAGTGTAAACA 602  
Db 481 ATCCCCCACCACCTGAGCAAGAGGTTGATCGGCCATGCTGTGAGCGAGTGTAAACA 540  
Qy 603 TTGCAACCCCGAGCAAGGAGCTGAAATCTCTGTGTAAACATCTGTGTGAGCTTCC 662  
Db 541 CTGCAACCCCGAGCAAGGAGCTGAAATCTCTGTGTAAACATCTGTGTGAGCTTCC 600  
Qy 663 ATTGACCCCGAGATGCTGAGAGGACTGAGCAACAGAGCTGAGTTTCCCTGCAATGC 722  
Db 601 ATTGACCCCGAGATGCTGAGAGGACTGAGCAACAGAGCTGAGTTTCCCTGCAATGC 660  
Qy 723 CATGTGACAGAGTGAACGCTTCACTACCCCTGCTTCCCAACAGTCCCTAACCCGGG 782  
Db 661 CBTGGAGAGAGTGAACGCTTCACTACCCCTGCTTCCCAACAGTCCCTAACCCAGCT 720  
Qy 783 CATGCCGCTGCGAGTGGTCTCTGCGGGGGAGAGCCGCACTCTGTGCTGAGCTCACCTTC 842  
Db 721 CATGCCGCTGCGAGTGGTCTCTGCGGGGGAGAGCCGCACTCTGTGCTGAGCTCACCTTC 780  
Qy 843 CGAAGCTTGTATGTCGCTCCCTGTGATGAGCATGAGCACTGTGACCTGTATGAT 902  
Db 781 CGAAGCTTGTATGTCGCTCCCTGTGATGAGCATGAGCACTGTGACCTGTATGAT 840  
Qy 903 AGCCTGAGCCCGCATGGAACCCCAAGCTGTGTGAGGCTGTGTGACCTTCTCACTCC 962  
Db 841 AGCCTGAGCCCGCATGGAACCCCAAGCTGTGTGAGGCTGTGTGACCTTCTCACTCC 900  
Qy 963 TACAACCTGACTTCTCTCTCTCCAGAAAGCTTCTCTGTGCAAGCTGATTAACCAAT 1022  
Db 901 TACAACCTGACTTCTCTCTCTCCAGAAAGCTTCTCTGTGCAAGCTGATTAACCAACT 960  
Qy 1023 GACCGGAGACATCTGAGCTTTGAGGCGACCTTCTTCCAGCTGCGCAAGATGAGAGCTGT 1082  
Db 961 GACCGGAGACATCTGAGCTTTGAGGCGACCTTCTTCCAGCTGCGCAAGATGAGAGCTGT 1020  
Qy 1083 GGGCGCTTTTGAATGACACCCCAAGGACATTTAGCAGCCCTTACTTCCAGGCTACTAC 1142  
Db 1021 GGGCGCTTTTGAATGAGACCCCAAGGACATTTAGCAGCCCTTACTTCCAGGCTACTAC 1080  
Qy 1143 CCGCCCAACATCAATGCAATGGAATATGAGAGTGGCCCAACACCGGAACGTGAAGTA 1202  
Db 1081 CCGCCCAACATCAATGCAATGGAATATGAGAGTGGCCCAACACCGGAACGTGAAGTA 1140  
Qy 1203 CGCTTCAAACTCTTCTATCTGTGTGAGCCCAACGTAACAGTGGCTCTTGACCAAGGAC 1262  
Db 1141 CGCTTCAAACTCTTCTATCTGTGTGAGCCCAACGTAACAGTGGCTCTTGACCAAGGAC 1200  
Qy 1263 TATGTGAGATCAACGAGGAGAACTGCTGGGTGAGAGTCCAGCTTTGTGTGAGCAGC 1322  
Db 1201 TATGTGAGATCAACGAGGAGAACTGCTGGGTGAGAGTCCAGCTTTGTGTGAGCAGC 1260  
Qy 1323 AACGACGCAAGATTAAGTCACTTCATCTGATCACTGTCACAGCAACCGGGTTC 1382  
Db 1261 AACGACGCAAGATTAAGTCACTTCATCTGATCACTGTCACAGCAACCGGGTTC 1320  
Qy 1383 CTAGCTAGTACTCTCTCTTCACTCAACGACCCGTCGAGGAGATTTATGTGCAAG 1442  
Db 1321 CTAGCTAGTACTCTCTCTTCACTCAACGACCCGTCGAGGAGATTTATGTGCAAG 1380  
Qy 1443 ACTGAGCGGTGATCCGAAGGAACTGGGCTGCGACCGCTTGGGCAAGATGCGCGGATAT 1502  
Db 1381 ACTGAGCGGTGATCCGAAGGAACTGGGCTGCGACCGCTTGGGCAAGATGCGCGGATAT 1440  
Qy 1503 AGTATGAGCGCTTCTGCGGATGCAATGCAACGACCCGTCGAGGAGATTTATGTGCAAG 1562  
Db 1441 AGTATGAGCGCTTCTGCGGATGCAATGCAACGACCCGTCGAGGAGATTTATGTGCAAG 1500  
Qy 1563 TCGAAGCCCTCTTCTGTGGTCTGTGACAGTGTCAAAGCTGTGGGAGCGAAGTGAAG 1622  
Db 1501 TCGAAGCCCTCTTCTGTGGTCTGTGACAGTGTCAAAGCTGTGGGAGCGAAGTGAAG 1560

QY 1623 GAGGGCTGACAGTCTGCTGCTGGAGTTTCAAGTTTCAATGGGAAGTGTCTCCCTCAG 1682  
 DB 1561 GAGGGCTGACAGTCTGCTGCTGGAGTTTCAAGTTTCAATGGGAAGTGTCTCCCTCAG 1620  
 QY 1683 AGCCAGAGTGTATGGGAAGACAACTGTGAGATGGGTCTGACGAGCTTCAATGAG 1742  
 DB 1621 AGCCAGAGTGTATGGGAAGACAACTGTGAGATGGGTCTGACGAGCTTCAATGAG 1680  
 QY 1743 AGGTGATGTGTCTCTTGCACCAATATACCTACCGCTGCCAATATGAGCTCTGTCTG 1802  
 DB 1681 AATGTAATGCTCTCTCTTGCACCAATATACCTACCGCTGCCAATATGAGCTCTGTCTG 1740  
 QY 1803 AGCAAGGCAACCTGATGATGTATGGGAAGACAACTGTGAGATGGGTCTGACGAGCTTCAATGAG 1862  
 DB 1741 AACAAGGCAACCTGATGATGTATGGGAAGACAACTGTGAGATGGGTCTGACGAGCTTCAATGAG 1800  
 QY 1863 AACTGTGATGTGTGGGTGAGCTCTTACCAACAGGCTGCGGTGGTGGTGGACAGAT 1922  
 DB 1801 AACTGTGATGTGTGGGTGAGCTCTTACCAACAGGCTGCGGTGGTGGTGGACAGAT 1860  
 QY 1923 GCGGAGAGAGGCGAGTGGCTCTGAGAGTGGAGCTTCAAGCTTGGGCGACAGGCTTGG 1982  
 DB 1861 GCGGAGAGAGGCGAGTGGCTCTGAGAGTGGAGCTTCAAGCTTGGGCGACAGGCTTGG 1920  
 QY 1983 TGTGGGGCTCGCTCATCTCTCTCTGAGCTGGGTCTCTGAGAGCTATGCTTTAGAT 2042  
 DB 1921 TGTGGGGCTCGCTCATCTCTCTCTGAGCTGGGTCTCTGAGAGCTATGCTTTAGAT 1980  
 QY 2043 GACAAATTTTCAAGTACTCAGACTTCAAGATGTGAGAGGCTTCTGAGGTCTGCTGGAG 2102  
 DB 1981 GAGACATTTTCAAGTACTCAGACTTCAAGATGTGAGAGGCTTCTGAGGTCTGCTGGAG 2040  
 QY 2103 CAGAGAGAGGAGTGGCTCTGAGAGTGGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2162  
 DB 2041 CAGAGAGAGGAGTGGCTCTGAGAGTGGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 2100  
 QY 2163 CCTCTCTTCAATGATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2222  
 DB 2101 CCTCTCTTCAATGATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2160  
 QY 2223 GTGAGTACAGAGCTGCTGCTGAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2282  
 DB 2161 GTGAGTACAGAGCTGCTGCTGAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2220  
 QY 2283 GCTGCAAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2342  
 DB 2221 GCTGCAAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2280  
 QY 2343 CTGATCTGCAAGAGGAGTATCGGTGATCAACCAAGCTTGTGAGAGCTTCACTTCACTTCACT 2402  
 DB 2281 CTGATCTGCAAGAGGAGTATCGGTGATCAACCAAGCTTGTGAGAGCTTCACTTCACTTCACT 2340  
 QY 2403 CCGAGCAGATCACCCACGAATGATGTGTGAGGTTTCTCAGTGGGGGTGTGAGCTCC 2462  
 DB 2341 CCGAGCAGATCACCCACGAATGATGTGTGAGGTTTCTCAGTGGGGGTGTGAGCTCC 2400  
 QY 2463 TGGCAGAGGCTGATCTGTGAGCTTGTGCAAGCGGAGAAAGATGGGCTGATGCTCCAG 2522  
 DB 2401 TGGCAGAGGCTGATCTGTGAGCTTGTGCAAGCGGAGAAAGATGGGCTGATGCTCCAG 2460  
 QY 2523 GCTGATGTGATGAGCTGGGGGTGAGAGGCTGCTCAGAGGAAACAGCCGCTGTACACA 2582  
 DB 2461 GCTGATGTGATGAGCTGGGGGTGAGAGGCTGCTCAGAGGAAACAGCCGCTGTACACA 2520  
 QY 2583 AGGCTCCTGTAGTGTGAGCTGATCAAGAGCACTGAGGCTATG 2630  
 DB 2521 AGGATCCTGTAGTGTGAGCTGATCAAGAGCACTGAGGCTATG 2568

RESULT 6  
 BC030532 3273 bp mRNA linear PRI 07-OCT-2003  
 LOCUS BC030532  
 DEFINITION Homo sapiens suppression of tumorigenicity 14 (colon carcinoma),

## FEATURES

source

## REMARK

COMMENT

## JOURNAL

JOURNAL

## AUTHORS

AUTHORS

## TITLE

TITLE

## JOURNAL

JOURNAL

## MEDLINE

MEDLINE

## PUBMED

PUBMED

## REFERENCE

REFERENCE

## JOURNAL

JOURNAL

## AUTHORS

AUTHORS

## TITLE

TITLE

## JOURNAL

JOURNAL

## MEDLINE

MEDLINE

## PUBMED

PUBMED

## REFERENCE

REFERENCE

## JOURNAL

JOURNAL

## AUTHORS

AUTHORS

## TITLE

TITLE

## JOURNAL

JOURNAL

## MEDLINE

MEDLINE

## PUBMED

PUBMED

## REFERENCE

REFERENCE

## JOURNAL

JOURNAL

## AUTHORS

AUTHORS

## TITLE

TITLE

## JOURNAL

JOURNAL

## MEDLINE

MEDLINE

## PUBMED

PUBMED

matrilase, epithelial), mRNA (cDNA clone MGC:40392 IMAGE:5213189), complete cds.  
 BC030532  
 ACCESSION BC030532.1 GI:20988874  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F., Stappleton, W., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 2 (bases 1 to 3273)  
 Strausberg, R.  
 Direct Submission  
 Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing By: The I.M.A.G.E. Consortium (LNL)  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nigri.nih.gov](mailto:nisc_mgc@nigri.nih.gov)  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, P., Latic, P., Legaspi, R., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Masello, C., Maekel, B., Mastrian, S.D., McCloskey, U.C., McDowell, J., Pearson, R., Stantryp, S., Thomas, P.J., Touchman, D.W., Tsurgon, C., Vogt, U.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAX Plate: 64 Row: K Column: 15  
 This clone was selected for full length sequencing because it passed the following selection criteria: similarity but not identity to protein.  
 Location/Qualifiers  
 1..3273  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:40392 IMAGE:5213189"  
 /tissue="Blood, adult leukocytes"



```

Db      1316  ACTACGTGAGATCAATGGGGAATAATCTGCGAGAGAGGTCCAGTCTGTCACCA 1375
Qy      1321  GCAACAGCAGAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGACACCGGCT 1380
Db      1376  GCAACAGCAGAGATTACAGTCTGCTTCCACTCAGATCACTGATCAACGACACCGGCT 1435
Qy      1381  TCCAGAGTGAATCTCTCTGAGCTCAACGACCCGTCGCGAGGATGTTCAATGACA 1440
Db      1436  TCTTAGTGAATCTCTCTGAGCTCAACGATCCATGCTCCGCGGAGTTCACGTGCC 1495
Qy      1441  AGACTGACGATGATCCGAAAGGAATGCGCTGCGAGCGCTGGGAGACTGCTCCGAT 1500
Db      1496  GCACGGGGGCGGTATTCGGAAGAGCTGCGCTGTGATGCTGCGGCTGACCGACCC 1555
Qy      1501  ATAGTATGAGCTTACTGCGGATGCAATGCCACCCACAGTTCACCTGCAAAAACAGT 1560
Db      1556  ACAGCATGAGCTCAACTGCAATGCGAGCGCGCCACCGATTCAGTGCAGAGACAGT 1615
Qy      1561  TCTGCAAGCCCCCTCTTCTGAGTCTGAGACAGTCTCAACGATGCGGAGAGAG 1620
Db      1616  TCTGCAAGCCCCCTCTTCTGAGTCTGAGACAGTCTGAGACAGTCTGAGACAG 1675
Qy      1621  AGAGAGGCTGACAGTCTGCTGAGGAGTTCAGGTTCAAGTTCATGGAAGTGTCTCTC 1680
Db      1676  AGCAGGGGTGACAGTCTGCTGAGGAGTTCAGGTTCAAGTTCATGGAAGTGTCTCTC 1735
Qy      1681  AGAGCAGAGTGTATGCGAAGACACACTGTGAGATGCGTCTGACAGGCTTCATG 1740
Db      1736  AAAGCCAGAGTGCATGGGAAGAGACGACTGTGGGAGCGGCTCCGACGAGGCTCTG 1795
Qy      1741  AAGAGTGAATGCTGCTCTTCTGACCAATAATACCTACCGCTGCGCAAAATGCGCT 1800
Db      1796  CCAAGGTGAAGCTGCTGCTCTTCTGACCAATAATACCTACCGCTGCGCTGCTGCT 1855
Qy      1801  TGAGCAAGGCGCAACCTGAGTGTGATGGAAGACGGAAGTGTGCTCCGATGAGA 1860
Db      1856  TGAGCAAGGCGCAACCTGAGTGTGATGGAAGACGGAAGTGTGCTCCGATGAGA 1915
Qy      1861  AAAAATGTAATGCTGCTGCTGCTGCTTACCAACAGAGCTGCGGTGTGTTGAGACA 1920
Db      1916  AGGACTGCACTGTGGGCTGCGGTATTCACGAGACAGGCTGTGTTGTTGGGGGACA 1975
Qy      1921  ATGCGGAGGAGGCGGAGTGGCTTGGGAGAGTGAAGCTTCCAGCTGCGGCGACCA 1980
Db      1976  ATGCGGAGGAGGCGGAGTGGCTTGGGAGAGTGAAGCTTCCAGCTGCGGCGACCA 2035
Qy      1981  TGTGTGGGGCTCGCTCATCTCTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db      2036  TGTGTGGGGCTCGCTCATCTCTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 2095
Qy      2041  ATGACAAAATTTCAGTACTCAGACTACAGCATGTGACGCTTCTGCTGCTGCTG 2100
Db      2096  ATGACAGAGATTCAGTACTCAGACTACAGCATGTGACGCTTCTGCTGCTGCTG 2155
Qy      2101  ACCAGAGAGGCGGAGTGGCTTGGGAGAGAGTGAAGCTCAACGATCATCACCC 2160
Db      2156  ACCAGAGAGGCGGAGTGGCTTGGGAGAGAGTGAAGCTCAACGATCATCACCC 2215
Qy      2161  ACCCTTCTCATATATTCAGCTTCACTGATGATGCTGCTGCTGCTGCTGCTGCT 2220
Db      2216  ACCCTTCTCATATATTCAGCTTCACTGATGATGCTGCTGCTGCTGCTGCTGCT 2275
Qy      2221  CGTGTGAGTACAGACCGTCTGCGGCCCATCTGCTGCTGCTGCTGCTGCTGCT 2280
Db      2276  CGGAGAGTACAGCTCCATGCTGCGGCCCATCTGCTGCTGCTGCTGCTGCTGCT 2335
Qy      2281  CTGCTGAGAGGCGATCTGGGTCAACAGCTGCGGCGACACAAAAGAGGAGTACCG 2340
Db      2336  CTGCTGAGAGGCGATCTGGGTCAACAGCTGCGGCGACACAAAAGAGGAGTACCG 2395
Qy      2341  CGCTGATCTCTGACAGAGGAGTGAATCGTGTATCATCAACCACTGTGAGAGCTCA 2400

```

```

Db      2396  CGCTGATCTGTCAAAAAGGATGAGATCCGCGTATCAACGACCACTCTGGAGACCTCC 2455
Qy      2401  TGCCGAGAGGATATCACCCGCAATGATGTGTGTGTTTCTTCACTGCGGAGTGTGACT 2460
Db      2456  TGCCGAGAGGATATCACCCGCAATGATGTGTGTGTTTCTTCACTGCGGAGTGTGACT 2515
Qy      2461  CTGCGAGGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db      2516  CTGCGAGGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2575
Qy      2521  AGGCTGTGTGTGATGAGCTGCGGAGTGAAGGCTGCGCTCAGAGAAACAGCGCTGTACA 2580
Db      2576  AGGCTGTGTGTGATGAGCTGCGGAGTGAAGGCTGCGCTCAGAGAAACAGCGCTGTACA 2635
Qy      2581  CAAGGCTCTCTGTGATGTTTGGGATGTGATCAAGAGACACTGCGGATATGACAGATGAC 2640
Db      2636  CAAGGCTCTCTGTGATGTTTGGGATGTGATCAAGAGACACTGCGGATATGACAGATGAC 2693
Qy      2641  AGACAGCCGACCAACCAACCCGACAGGATGCGCGACATGACACCTGTATACAGAGAG 2700
Db      2694  -GCCACCAAAATGTGTACACTGCGGAGGCGACCACTATGTCACCCAGTGTGACG- CCT 2751
Qy      2701  GAACTGACGACATTTATGCTGTGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2760
Db      2752  GCAGGCTGAGACTGTGACCGCTGACTGTGACCAAGCGCC- CAGAAATATCATGTGACT 2810
Qy      2761  GCATCTTATGAGCTGAGCT 2780
Db      2811  CAATCTCCAGGCTCCCAAT 2830

```

```

RESULT 7
BD274671 3149 bp DNA linear PART 17-JUL-2003
LOCUS BD274671
DEFINITION Matrixphage, serine protease and application of the same.
ACCESSION BD274671.1 GI:33084439
VERSION JP 2002539093-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3149)
AUTHORS Dickinson, R. B., Lin, C., Johnson, M., Wang, S., and Nyne, I.
TITLE Matrixphage, serine protease and application of the same
JOURNAL Patent: JP 2002539093-A 2 19-NOV-2002;
GEOGROWTH UNIVERSITY
OS Homo sapiens (human)
PN JP 2002539093-A/2
PD 19-NOV-2002
PF 10-MAR-2000 JP 2000603721
PR 12-MAR-1999 US 60/124006
PI ROBERT B DICKSON, CHEN-YONG LIN, MICHAEL JOHNSON, SHAOWENG WANG,
PI ISTVAN ENYEDI
PC A61K45/00, A61K35/78, A61K51/00, A61P17/00, A61P35/00, PC
A61P43/00,
PC A61P43/00, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12N9/64, C12N15/09,
PC C12Q1/37, GOIN33/53, GOIN33/577//C12P21/08, A61K49/02, C12N15/00,
PC C12N5/00
CC Matrixphage, serine protease and application of the same FH
KEY Location/Qualifiers
FT source 1..3149
FT /organism="Homo sapiens (human)".

```

## FEATURES

source

Location/Qualifiers  
1..3149  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 60.4%; Score 1889.8; DB 6; Length 3149;  
Best Local Similarity 81.4%; Pred. No. 0;

Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;				
QY	50	GA	CGCGAAACCAATGGTAGAATGGGGCCGCAAGCCCGAGGGGGCTTCAAGACTT	109
Db	23	GG	CTCGGGGACCAATGGGAGGATGGGGCCGCAAGGGCGGGGGCCGAAAGACTT	82
QY	110	CG	CGCGGGGACCTCAATCAACAATCCCGGGCTAGAGAACTAATAGCTTTGAGGAGGGTGT	169
Db	83	CG	CGCGGGGACCTCAATCAACAATCCCGGGACAGAGAAATGAGCTTTGAGGAGGGTGT	142
QY	170	GG	AGTTCTGCTCGACCAATGCAAGAAAGTGAAGAACGAGCCCAAGCGCTGGT	229
Db	143	GG	AGTTCTGCTCGACCAATGCAAGAAAGTGAAGAACGAGCCCGGGGCGCTGGT	202
QY	230	GG	TGCTGGTGGAGTGTCTTCACTTCTTCTGCTCTCCCTCATGGCTGGCTGGT	289
Db	203	GG	TGCTGGTGGAGTGTCTTCACTTCTTCTGCTCTCCCTCATGGCTGGCTGGT	262
QY	290	GT	GGCACTTCCATTCATGCAATGCGGGTTCAAAAAGTTCATGAGCCATCTGAGAT	349
Db	263	GT	GGCACTTTCAGATACCGGAGCTGCTGTTCAGAGAGCTTCATATGCAATGAGAT	322
QY	350	CA	CAAAATGAGATCTTCTGATGCGTATGAGAACTCCACTCAAGATTATACGCT	409
Db	323	CA	CAAAATGAGAAATTTGTGATGCTTACGAGAACTCAACTCCACTGAGTTTGAAGCT	382
QY	410	GG	CGAGCAGGTGAAGAGCGCTGAAGCTGTGTAAATGAAGTCCCTGGCTGGCTC	469
Db	383	GG	CGAGCAGAGGTGAAGAGCGCTGAAGCTGTGTAAAGTCCCAATTCCTGGGCTC	442
QY	470	CT	CAACAAGAGCGGCTTAATCTGCTTCACTGAGAGGAGTGTCAATGCTTACTAG	529
Db	443	CT	CAACAAGAGAGTGGCTGTGAGAGGCTTCAAGAGGAGGAGGAGGCTCAATGCTTACTAG	502
QY	530	GT	CAGAGTTGAGATCCCGCCCAACTGTGCAAGAAAGAGTTGATGCGCATGCTGTGA	589
Db	503	GT	CAGAGTTGAGATCCCGCCAGCACTGTGAGAGAGAGCGCGCTCATGAGCGAGGA	562
QY	590	GG	CGATGTGAATGCTTGCACCCCGAGAGCGGGCACTGAAATCCTTGGTCAATCATCTGT	649
Db	563	GG	CGATGTGAATGCTTGCACCCCGAGAGCGGGCACTGAAATCCTTGGTCAATCATCTGT	622
QY	650	GG	TGCTGCTTCCCAATTTGACCCCAAGATGCTCAGAGAGACTGAGAACACAGCTGCACTT	709
Db	623	GG	TGCTGCTTCCCAATTTGACCCCAAGATGCTCAGAGAGACTGAGAACACAGCTGCACTT	682
QY	710	TG	CCCTGATGCCATGGTGTGAGAGAGTGAACGCTTCACTAACCTTGGCTTCCCAAG	769
Db	683	TG	CCCTGATGCCATGGCTGTGAGAGTGTGATGCTTCACTAACGCTTCCCTGAG	742
QY	770	TC	CCATCCCGGCGATGCGCGCTGCTCAATGGGTCTGTGGGGGGGAGCGCGACTGTGGCT	829
Db	743	CC	CCATCCCGGCGATGCGCGCTGCTCAATGGGTCTGTGGGGGGGAGCGCGACTGTGGCT	802
QY	830	GAG	CTCACTTCCGAAGCTTTGATGTGCTCCCTGTGATGAGATGAGATGAGCTTGGT	889
Db	803	GAG	CTCACTTCCGAAGCTTTGATGTGCTCCCTGTGATGAGATGAGATGAGCTTGGT	862
QY	890	CA	CCGTATGATAGCTGAGCGCCCATGAGAACCCCAAGCTGTGTGGCTGTGTGGAC	949
Db	863	CA	CCGTATGATAGCTGAGCGCCCATGAGAACCCCAAGCTGTGTGGCTGTGTGGAC	922
QY	950	CT	TTCATCCCTCTCTCAACCTTGAATCTTCTCTCCCAAGAGCTTCTTCTTCAAGCT	1009
Db	923	CT	TTCATCCCTCTCTCAACCTTGAATCTTCTCTCCCAAGAGCTTCTTCTTCAAGCT	982
QY	1010	GATA	CAATACTGACCGGCGACATCTGGCTTTGAGGCACTTTCTTCACTGAGCCCA	1069
Db	983	GATA	CAATACTGACCGGCGGCACTTCCGGCTTTGAGGCACTTCTTCACTGAGCCCA	1042
QY	1070	GAT	GAGCACTGTGGCGCTTTTGAAGTGAACCCAGGAGCACTTTAGAGCCCTCACTA	1129
Db	1043	GAT	GAGCACTGTGGAGCGCTTGAAGTGAAGCCAGGAGCACTTTAGAGCCCTCACTA	1102

QY	1130	TC	AGGCGCACTACCGCGCCCAATCAATGCACTGCAATGAAATATCAAGTGTCCCAACAGG	1189
Db	1103	CC	AGGCGCACTACCGCAATCAATGCACTGCAATGAAATATGAGTGTCCCAACAGCA	1162
QY	1190	GA	AGTGAAGGTGGCTTTAACTCTTATCTGTGGGAGCCCAAGTACAGTGGGCTC	1249
Db	1163	GA	AGTGAAGGTGGCTTTAAATCTTCTTACCTGTGAGCCCGGCTGTCTGGGGCAC	1222
QY	1250	CT	GACCAAGACTATGTGAGATCAACGGGGAGAGTACTGCGGTGAGAGTGTCCAGTT	1309
Db	1223	CT	GACCAAGAGACTATGTGAGATCAATGGGAGAAATCTGTGAGAGAGTGTCCAGTT	1282
QY	1310	TGT	GAGCAGCAAGCAAGCAAGATTAAGTCCACTTTCATTTCTGATATCTCTGATAC	1369
Db	1283	CG	TGTCAACAGCAAGCAAGATCAAGTTCAGTTCCTTCCACTCAAGTATCACTTAC	1342
QY	1370	GA	ACCGGGTCTCTGAGTGAAGTCTCTCAAGCTCCCAAGACCGCGGCCAGAGAT	1429
Db	1343	GA	ACCGGGTCTCTGAGTGAAGTCTCTCAAGCTCCCAAGACCGCGGCCAGAGAT	1402
QY	1430	GT	CATGTGCAAGACTGCAAGCTGCAATCCGAAAGAACTGCGCTGCGACGCTGGGCA	1489
Db	1403	GT	CATGTGCGGCAAGCGGGGCTGTATCCGAAAGAGCTGCGCTGTGATGCTGGGCGA	1462
QY	1490	CT	CGGATTAATATGATGAGGCTTATCTGCCATGCAATGCCACCAACCAATTAAGTG	1549
Db	1463	CT	CGACCGGATTAATGATGAGGCTTATCTGCCATGCAATGCCACCAACCAATTAAGTG	1522
QY	1550	CA	AAACCAATTTGCAAGCGCCCTCTTCTGAGTCTGTGACAGTGTCAACAGACTGTGGGA	1609
Db	1523	CA	AAACCAATTTGCAAGCGCCCTCTTCTGAGTCTGTGACAGTGTGAACAGACTGTGGGA	1582
QY	1610	CG	AGTGAAGAGGAGGCTGCAAGCTGCTCTGAGAGTTTCAAGTGTTCATATGGGAA	1669
Db	1583	CA	AGAGAGAGAGGAGGCTGCAAGTGTCTCGGCCCAACCTTCAAGTGTTCATATGGGAA	1642
QY	1670	GT	TCTCTCCTCAGAGCCCAAGTGTAAATGGGAAAGAACTGTGGAATGGGCTGTGAGA	1729
Db	1643	GT	TCTCTCCTCAGAAAGCTGCAAGTGTAAATGGGAAAGAACTGTGGAATGGGCTGTGAGA	1702
QY	1730	GG	CTCATATGAGACGCGTGAATGTCTCTCTTGACCAAAATATATCTAACGCTGCCAAA	1789
Db	1703	GG	CTCTCTCCTCCCAAGTGAACCTGTCTACCTTGTACCAAAACACTAACGCTGCCAAA	1762
QY	1790	TG	GCCTGTGTGAGCAAGGCAACCTGATGTGATGGGAAAGCACTGTAGCGATGG	1849
Db	1763	TG	GCCTGTGTGAGCAAGGCAACCTGATGTGAGCGGAAAGGACTGTAGCGACGG	1822
QY	1850	CT	CGATGAGAAAACTGTGACTGTGGCTGTGCAATCTTTTACAAACAGGCTGGGTGT	1909
Db	1823	CT	CGATGAGAAAGACTGTGCACTGTGGCTGTGCAATCTTTTACAAACAGGCTGGGTGT	1882
QY	1910	TG	GGGCAAGATGCGAGCGAGGAGTGGCTCTTGAGAGGTGAGCTTCAAGCCCTGGG	1969
Db	1883	TG	GGGCAAGATGCGAGTGTGAGGAGTGGCTCTTGAGAGGTGAGCTTCAAGCTCTGGG	1942
QY	1970	CC	AGGCACTTGTGTGGGGCTGTGATCTTCTCTGACTGTGTCTGTGACGCTCA	2029
Db	1943	CC	AGGCACTATGTGGGTCTTCTCTCATCTTCTTCAACTGAGCTGTGTCTGTGCGACA	2002
QY	2030	TT	GCTTTAGAGATGCAAAAATTTTCAAGTACTCAAGACTCAAGATGTGAGCGGCTTCT	2089
Db	2003	CT	GCTTACATGATGCAAGAGATTTCAAGTACTCAAGCCCAAGCAAGTGTGAGCGGCTTCT	2062
QY	2090	GG	GTGTGTGAGCAAGACGAGTGTCTTGGGGTGTGAGAGCTGAAGCTCAAAAG	2149
Db	2063	GG	GTGTGTGAGCAAGACGAGCGAGGCTTGGGGTGTGAGAGCTGAAGCTCAAAAG	2122
QY	2150	TAT	CAATCAACCACTTCTCTTCAATGATTTCACTTTGACATATGACTGTGCTGTGA	2209
Db	2123	CAT	CAATCTTCCAGCCCTTCTTCAATGATTTCACTTTGACATATGACTGTGCTGTGA	2182







```

RESULT 9
AF133086 3120 bp mRNA linear PRI 01-OCT-1999
LOCUS Homo sapiens membrane-type serine protease 1 mRNA, complete cds.
DEFINITION AF133086
ACCESSION AF133086
VERSION AF133086.1 GI:6002713
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Takeuchi, T., Shuman, M.A. and Craik, C.S.
TITLE Reverse biochemistry: use of macromolecular protease inhibitors to
dissect complex biological processes and identify a membrane-type
serine protease in epithelial cancer and normal tissue
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11054-11061 (1999)
MEDLINE 99432178
PUBMED 10500122
REFERENCE
AUTHORS Takeuchi, T., Shuman, M.A. and Craik, C.S.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-1999) Dept. Pharm. Chem., University of
California, San Francisco, 513 Parnassus Ave., Box 0446, San
Francisco, CA 94143-0446, USA
FEATURES
source Location/Qualifiers
1..3120
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="PC-3"
/cell_type="prostate carcinoma"
37..7604
/note="Mt-Sp1; chymotrypsin clan serine protease"
/codon_start=1
/product="membrane-type serine protease 1"
/db_xref="GI:6002714"
/translation="MGSDBARKGGGPKDFGALKYNSHEKNGLEGEVPLPVNVV
KVEKHGPERWVLAVALIGLLVLLGIFLWMLQYRVRVQKFNQMTLNENFV
DAVENSNGRWFVLAVALIGLLVLLGIFLWMLQYRVRVQKFNQMTLNENFV
IPOHVEAEERVAERVYMLPPRARSLSFVTSVNVKTVQVONQONSCSGL
HARGVLEMTTGPFPDSDPYPAHARCOMALRGDADSVALSTRSDLSADERSGLV
TYVNTLSMEPMALVOLGCTYPPSYNLTTHSQQVLLTLTNTERRRHGFATFOL
PRMSCGGRLEKRAQGTENSPLYGHYPNIDCTANI EYNNQHVVRKFEFLIEPGV
PAGTCPKDYVEINTEGRKYGERSQFVTSNSKITYRPHSDQYDTGTFLAEYLSDDS
DPCPGQTCRTGRCIRKELRCGMADCTHDELNCSQDAGQPTCKXKPKPLFWVC
DSVNDGNSDEQSCPAQCFRCSNGKCLSKSQCNKGDGDSDEASCPKXVVT
CTKHTRCLNGCLSKNGPECDGKEDSGDEKDCDGLRFTFOAVVGSTDADEG
EMPQVSLHALGGQHCICGASLISPMWLVSAHCYIDDKGFRSDTPTWATFGLHDQS
ORSAFQVERLRKRLIISHPFENDFTFDYDIALLELEKPAEYSSMVRPCLPDASHVFP
AGKAIWVGWHTYGGTGALILQGEIRVNOTTCENLIPQITPRMWCFTSLGSGV
DSCGDSGGPLSSVEADGRIFQAGVAVSWGDGCAQRNKPQVYTRPLPLFRDWKENTGV"
ORIGIN
Query Match 60.7%; Score 1884.8; DB 9; Length 3120;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 447; Indels 1; Gaps 1;

```

```

Db 188 GCCCGGGGCGCTGGGTGCTGAGCCGCTGATCGGCTCTTGTTGCTGG 247
Qy 274 TGCTGGCTTGCTGTTGTCACCTTCATATAGGAATGGCGGTTCAAAAAGTCTTCA 333
Db 248 GGATCGGCTTCTGTTGTCATTTGAGTACCGGAGCGTGGCTTCAAGAGTCTTCA 307
Qy 334 ATGGCCATCTGAGANTACAAATAGATCTTTCGTGATGCTATGAGACTCCACTCCA 393
Db 308 ATGGCTACATGAGATACAAATGAGAAATTTTGTGATGCTATGAGAACTCCACTCCA 367
Qy 394 CAGAGTTTATACCTGCGCAGCCAGGTGAAGAGCGCTGAAGCTGCTGTACANTAG 453
Db 368 CTGAGTTTGAACCTGGCCAGAGGTGAAGAGCGGCTGAAGCTGCTGTACAGCGGAG 427
Qy 454 TCCCTGCTGGGCTCCCTACCAAGAGTGGCTGTATCTGCTTCACTGAGGCGAGTG 513
Db 428 TCCCATTTCTGGGCCCCCTACCAAGAGTGGCTGTGAGGCGCTTCAAGAGGCGAGC 487
Qy 514 TCATGCTACTACTGCTGAGTTCAGATCCGCCACACCTGGCGAGAAAGGTTGATC 573
Db 488 TCATGCTACTACTGCTGAGTTCAGATCCGCCAGACCTGGTGGAGAGGCGAGC 547
Qy 574 GCGCCATGCTGTGAGCGAGATTGAACATTTGCCACCCGAGCAGCGGCACTGAAATCCT 633
Db 548 GCGTCATGGCCAGAGCGCGTATGATCTGCTGCCCGCGGCGGCTCCCTGAAGTCT 607
Qy 634 TCGTGTAAATCTGTTGGCTTCCCATTTGACCCCGAAGATCTGCAAGAGATCAG 633
Db 608 TTGTGTACCTCAGTGTGGCTTTCCACGAGATCCAAACAGTACAGAGAGCCAGG 667
Qy 694 ACAACAGCTGCAAGTTTGGCTGCAATGCGCATGTCAGAGTTCAGTCAATCC 753
Db 668 ACAACAGCTGCAAGTTTGGCTGCAATGCGCATGTCAGAGTTCAGTCAATCC 727
Qy 754 CTGGTCTCCCAACAGTCCCTTACCGGCGCATGTCGCTGCAAGTGGTCTTGCGGAG 813
Db 728 CGGGCTTCCCTACAGCCCTTACCGGCGCATGTCGCTGCAAGTGGTCTTGCGGAG 787
Qy 814 AGCGGAGCTGTGCTGAGCCCTACCTTCCGAGCTTGTATGCTCTCTCTGATGAGC 873
Db 788 AGCGGAGCTGTGCTGAGCCCTACCTTCCGAGCTTGTATGCTCTCTCTGATGAGC 847
Qy 874 ATGGCAGTACCTGTGACCTGTATGATAGCTGAGCCCATGGAACCCACGCTGTGG 933
Db 848 GGGGAGGAGCCTGTGAGCGGTGACCAACCTTGAAGCCCATGAGCCACGCTGTGG 907
Qy 934 TGGCGCTGTGTGACCTTCTTCACTCTTCAACCTTGAATCTTCTCTTCCAGAAC 993
Db 908 TGCAGTTGTGTGACCTTCTTCACTCTTCAACCTTGAATCTTCTCTTCCAGAAC 967
Qy 994 TCTTCTTGTACAGCTGATTAACCATTAAGTACCGGCGCATCTCTGCTTTGAGGCACTT 1053
Db 968 TCTTCTTGTACAGCTGATTAACCATTAAGTACCGGCGCATCTCTGCTTTGAGGCACTT 1027
Qy 1054 TCTTCAAGCTGCCAAGATGAGCACTGTGCGGCTTTTGTAGTACACCCAGGAGCAT 1113
Db 1028 TCTTCAAGCTGCCAAGATGAGCACTGTGCGGCTTTTGTAGTACACCCAGGAGCAT 1087
Qy 1114 TTAGAGGCTTATCTATCAAGGCACTTCCGCGCAATCACTCACTGCAATGAAATATCA 1173
Db 1088 TCAAGAGCTTATCTATCAAGGCACTTCCGCGCAATCACTGCAATGAAATATCA 1147
Qy 1174 AGTGGCCCAACACCGGAGGTGAAGGAGGCTTCAACTTCTGATGCTGAGCCCA 1233
Db 1148 AGTGGCCCAACACCGGAGGTGAAGGAGGCTTCAACTTCTGATGCTGAGCCCA 1207
Qy 1234 ACGTACAGTGGGCTCTTGCACCAAGACTTATGTGAGATCAAGGAGGAGTACTGCG 1293
Db 1208 GCGTGGCTGGGAGCACTGCGCCCAAGGACTTACGTGAGATCAATGGGAGGAAATCTGCG 1267
Qy 1294 GTGAGAGTCCCAATTTGTGTGAGCAGCAACAGCAGAGATTAAGTCCACTTCCATT 1353
Db 1268 GAGAGAGTCCCAATTTGTGTGAGCAGCAACAGCAGAGATTAAGTCCACTTCCACT 1327

```

QY	1354	CTGATCACTGCTACAGGACACCGGGTCTCTGCTAGTACACTCTCTACGACTCCAAAG	1413
Db	1328	CAGATCAAGTCTACACCGACACCGGCTTCTTACTGATAAACCCTCTCTACACTCCAGTG	1387
QY	1414	ACCCGTGCCCGAGGAGTATTCTATGTGCAGACCTTGACGCTGCATCCGAAAGAACCTGCGCT	1473
QY	1388	ACCCATATGCCCGGGGCAATTCACTGCGCGACGCGGGCGGTGTATCCGGAAGGACCTGCGCT	1447
QY	1474	GGACCGCTGGGGACGAGCTGCCGGATTTATGTATGAGGGTTATCTGCGCATGATGCCA	1533
Db	1448	GTAATGCTGGGCGCACTGACCGACCAACGCGATTAAGCTCAACTGCAAGTTGGACGCGG	1507
QY	1534	CCCAACGATTCAAGTGCAGAAAACAGTTCTGCAAGCCCTCTTCTGGTCTGTGACAGTG	1593
Db	1508	GCCACCAAGTTCAAGTGCAGAACAGATTGTGCAAGCCCTCTTCTGGTCTGTGACAGTG	1567
QY	1594	TCAACGACTGTGGGGACGGAAGTGCAGAGAGGGCTGCACTGTCCTGCTGGGAGTTTCA	1653
Db	1568	TGAACGACTCGGAGACAAACAGGACCAAGAGGGGTGCAGTTTCCGGCCACAGCCTTTCA	1627
QY	1654	AGATGTTCAATGGGAAGTGTCTCCCTCAGACCCAGACGATTAATGGGAAGACAACTGTG	1713
Db	1628	GGTCTTCCAAATGGGAAGTGTCTCTGAAAGCCAGACGATGCATATGGGAAGAGCACTGTG	1687
QY	1714	GAGATGGGCTTGACGAGGCTTTCATGTGACAGCGTGAAATGTCTCTCTTGACCAAAATATA	1773
Db	1688	GGGACGGGCTCCGACGAGGCTCTCCGCCCAAGGTGAACGTGTCACTTTGACCAAAACACA	1747
QY	1774	CTTACCGCTGCGCAAAATGGCCTCTGTCTGAGCAAGGCAACCTGAGTGTGATGGGAAGA	1833
Db	1748	CTTACCGCTGCTCTAATGGGCTCTGCTTGTAGCAAGGCAACCTGAGTGTGAGCGGAAGG	1807
QY	1834	CGGACTGTAGCGATGGCTCCGATGAGAAAACTGTGACTGTGGGCTGCGATCCTTTACCA	1893
Db	1808	AGGACTGTAGCGACGGCTCAAGATAGAAAGACCTGCACTGTGGGCTGCGGTCTCATTTACGA	1867
QY	1894	AACAGGCTTCGGTGTGTGTGTGGCAACGAATGGACGAGGGCGAGTGGCCCTGGGAGGTGA	1953
Db	1868	GACAGGCTCTGTTGTGTGGGGGCAACGAATGGGATGAGGGCGAGTGGCCCTGGCAGGTAA	1927
QY	1954	GCTTCACAGCCTTGCGGCGACAGGACCACCTGTGTGGGGCTCGCTATCTCTTGACTGGC	2013
Db	1928	GCTTGCAATGCTCTGGGCGAAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCAACTGGC	1987
QY	2014	TGTTCTCTGACGCTCATCTTCTTCAAGATGACAAAATTTCAATCACTATCAGACTACACA	2073
Db	1988	TGTTCTCTGCGCGCACACTGCTTACATCGATGACAGGATTTCAAGTACTCAGACCCACGCG	2047
QY	2074	TGTGGACGGCCCTTCTGCTGGTCTGTGACCAAGACCAAGCGAGTGCCTCTGGGGTGCAGG	2133
Db	2048	AGTGGACGGCTTCTTGGGCTTTGCACAGACCAAGACGCGGAGCCCTGTGGGGTGCAGG	2107
QY	2134	AGCTGAAGCTCAAAAGTATCATCAACCAACCTTCTTCAATGATTTCACTTTCGACTATG	2193
Db	2108	AGCGGAGGCTCAAGGCGCATCATCTCCACCCCTTCTTCAAGTCACTTCACTTTCGACTATG	2167
QY	2194	ACATGGCTTGCTGAGCTGGAGGAAGTGGGTGGAGTACAGACACGTCGTGGGCCCACTCT	2253
Db	2168	ACATGGCTTGCTGAGGAGCTGGAGGAACCGGACAGATTAAGCTTCATGATGTGGGCGCACTCT	2227
QY	2254	GCTTCTGTATGCTTACCATATGTCTTCCCTGCTGGCAAGGCCATCTGGGTACAAGCTGGG	2313
Db	2228	GCTTGGCGGAGCGCTCCCATATGTCTTCCCTGCGGCAAGGCCATCTGGGTACAAGCTGGG	2287
QY	2314	GGCACAACAAAAGGAGGTACCGGAGGCGGTATCTCTGACAGAGGGGTGAGATCGGTGCA	2373
Db	2288	GACACACCAAGTATGAGGCACTGGGCGGCTGATCTCTGAAAAGGGTATGATCCGCTGCA	2347
QY	2374	TCAACCAAGCAACTGTGTAGACTCATATCCGCAAGATGACCCCAAGATGATGTGTG	2433
Db	2348	TCAACCAAGCAACTGTGAGGAACCTCTTCCGCAACAGATACGCTGGCGCATGATGTGCG	2407

QY	2434	TGGGTTTCTCAGTGGGGGTGTGAGACTCTCTGCCAGGTGACACTGTGTCGCCCTTGTCAA	2499
Db	2408	TGGGCTTCTTCAGCGCGCGCGGTGAGACTCTCTTCCAGAGGTATTTCCGCGGAGCCCTGTGCCA	2467
QY	2494	GGCGGAGGAAAGATGGGCGCAATGTTCCAGAGCTGTGTGTGACTGTGGGTGAGAGCTGCG	2553
Db	2468	GGCTGAGGCGGATGGGCGGAGTTTCCAGGCCGCTGTGTGTGAGCTGGGAGAGAGGCTGCG	2527
QY	2554	CTCAGAGGAAACAAGCCAGGCGGTGTACACAAAGCTCCTTGAATTGGGAGCTGATCAAG	2613
Db	2528	CTCAGAGGAAACAAGCCAGGCGGTGTACACAAAGGCTCCTCTGTTCGGGAGCTGATCAAG	2587
QY	2614	AGCACTGGGGGTATATGACGATGAGACAACA	2645
Db	2588	AGAACACTGGGGGTATATGAGGCGCGGACACCA	2619
RESULT	10		
LOCUS	AB030036	3128 bp	mRNA
DEFINITION	Homo sapiens mRNA for prostamin, complete cds.		
ACCESSION	AB030036		
VERSION	AB030036.1	GI:12249014	
KEYWORDS	prostamin.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (sites)		
AUTHORS	Yamaguchi, N. and Mitsui, S.		
TITLE	Molecular cloning of a novel transmembrane serine protease expressed in human prostate		
JOURNAL	2 (bases 1 to 3128)		
AUTHORS	Yamaguchi, N. and Mitsui, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUL-1999) Nozomi Yamaguchi, Kyoto Prefectural University of Medicine, Res. Ins. Geriatrics, Kawatamachi Hirokoji, Kyoto, Kyoto 602-8566, Japan (E-mail: nozomi@koto.kpu-m.ac.jp, Tel:81-75-251-5848, Fax:81-75-251-5848)		
FEATURES	location/Qualifiers		
Source	1..3128		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="prostate"		
	16..2583		
	/codon_start=1		
	/product="Prostamin"		
	/protein_id="BAB20376.1"		
	/db_xref="GI:12249015"		
	/translation="MGSDPRARKGGGPKDFGAGLHKYRREKYNGLGEEVFLPVNVKVKRKHGRVVLAAVLIGLLVLLIGFLVHLQYRDVQVCFENGVMRTINENVDAYENSTEFVSLASKYKAKLILYGVLPGLGHHESAVTAAESGVLYAYSESPILQHLVEAEKVMKEERVMLPPARSLKSPVYTSVAFLPDSCTORTDNCSFGLHARVGEIMRFTTPGFPDSPPYPAHRCQWALRGDVSUSLITFRSFDLASCDBRSGLDLYTVNLTSPMEBHVLVOLCGTYPSPYNLTFRSSQNVLLITLTNTERHPEGEATFOLPRMSCGGRLLKAGCTGNSPYYPGHYPNIDCTNNIEVPNNQHKVFKFETLLEPGVPAGRCQDYVETNGEKYCGRSRCPVMTNSNKLTVRPHSDSYDTGFLAEVLYSDSSDPCHQPTCTRTGRCIRKEIRCDGADCTDHSDELNGSCDAGHPTCKNRCPLFWCDSVNDGNSDEGSCSPAQTFRCSNKKCLSKSQCKGKDDCGGSDSAQPKNVVLTCTKTRCLNGLCISKNGECDEKSDSDGDERDCDGLRFTROARVVGDTADGDEKEMTQVSLHAGGCHI CASGLSPNMLVSAHCYIDRGFRYSPTQWTVFLGLHSDOSORSAAPGVQERLKRILSHFPNDFTPVDILALELKEPAEYSSMWRICLPDASHVPEPAGKAIWTGKMHTOYCGTGALLDKGELRVINOTTCNLLPEQTLTPMNCVGFSGGVDSQGDSDGFLSVBADGRITQAGVSMGDCAGRNKRGVYTRLLPRDWIKENTGV"		
ORIGIN			
Query Match	60.7%	Score 1884.6;	DB 9; Length 3128;
Best Local Similarity	81.4%	Pred. No. 0;	
Matches 2222;	Conservative	0; Mismatches 504;	Indels 5; Gaps 3;

```
Db      3  GGCCTCGGGGACATGGGGAGCGATCGGGCCCCGAAAGGCGAGGGGGCCCGAAGAGCTT 62
Qy      110 CGGCGCGGACTCAAGTACAACTCCCGCTAGAGAACATGAATGGCTTTGAGAGGGTGT 169
Db      63  CGGCGCGGACTCAAGTACAACTCCCGCAAGAGAGTGAATGGCTTTGAGAGAGCGCT 122
Qy      170  GGAATTCCTGCTCGCAACAATGCGCAAGAAAGTGAAGAGCGAGGCCCGAAGCGCTGGT 229
Db      123  GGAATTCCTGCTCGCAACAACGTCGAAGAGTGAAGAGATGGCCCGGAGCGCTGGT 182
Qy      230  GGTGCTGATGGAGTGTCTGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 289
Db      183  GGTGCTGATGGAGCGGTGTGATCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
Qy      290  GTGGCACTTCATTATCGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCGAGAT 349
Db      243  GTGGCACTTCAGTATCGAGACGGGTGTGCGAAGAGTCTTCAATGGCTACATGAAGAT 302
Qy      350  CAGAAATGAGATCTTTTGTGATGCTATGAACTCCACTCCACAGATTATCAAGCT 409
Db      303  CAGAAATGAGAAATTTGTGATGCTATCGAAGATCTCCACTCCACTGATTTGTAAGCT 362
Qy      410  GGCACCCAGTGAAGAGGCGCTGAAGCTGTGTAACATGAAGTCCCTGTCTGTGGTTC 469
Db      363  GGCACCAAGTGAAGAGACCGGCTGAAGCTGTGTAACAGGAGTCCATTCCTGGGCTC 422
Qy      470  CTACCAAGAAAGTGGCTGTAACTGCTCTCACTGAGGAGTGTATCGCTACTACTG 529
Db      423  CTACCAAGAAAGTGGCTGTGAGCGGCTTCAAGAGGAGAGTGTATCGCTACTACTG 482
Qy      530  GTACAGGTTCAGATATCCCCCACTGCGCAGAGAGTGTATGCGGCAATGGCTGTGA 589
Db      483  GTCTGAGTTCAGATATCCCCCACTGCGCAGAGAGTGTATGCGGCAATGGCTGTGA 542
Qy      590  GCGAGTTGTAACATTTSCACCCCGACAGCGGCACTGAATCCCTGTGTAACTGTGT 649
Db      543  GCGGCTGAGTTCAGTGTGCTGCTGCGGAGCGGCTCTCTGAACTCTTGTGTGTCACTCA 602
Qy      650  GGTGCTTTCCTCCATTTGACCCCAAGATGTGCAAGAGACTCAAGACAGCTGACGTT 709
Db      603  GGTGCTTTCCTCCACGAGCTCCCAAAACAGTACAGAGAGCCAGAGACAGCTGACGTT 662
Qy      710  TGGCCCTGCAATGCGCATGTGTGAGAGTGAACAGCTTCACTACCTCTGCTCCCAAC 769
Db      663  TGGCCCTGCAATGCGCGGTGTGAGAGTGTGAGCTTCACTACCGCTCTCTCTGACAG 722
Qy      770  TCCCTACCCGCGCATGCGGCTGACAGTGGGTCTGCGGGGGAGACCGCACTTGTGT 829
Db      723  CCCCTACCCGCTCATGCTGCTGAGTGGGCTTGCAGGGGGAGACCGCACTCACTGT 782
Qy      830  GAGCTTCACTTCCGAAGCTTTGATGTGCTCTCTGTGATGAGCATGAGCATGACTGGT 889
Db      783  GAGCTTCACTTCCGAGCTTTGACCTTGCTGCTGTGAGCGAGCCGAGAGGACTGGT 842
Qy      890  CACCGTGAATGATTAACCTGAGAGCCCATGAGACCCCACTGTGTGTGTGTGTGAGAC 949
Db      843  GACGCTGAACACCTGAGAGCCCATGAGAGCCCATGAGAGCCCATGAGTGTGTGTGAGAC 902
Qy      950  CTTCTCACTCTCTCAACTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1009
Db      903  CTTACCTCTCTCTCTCAACTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 962
Qy      1010  GATTAACCAATATCTGACCGGCGACATCTGAGCTTTGAGGCACTTTCTTCAAGTCCCAA 1069
Db      963  GATTAACCAACATCTGAGCGGCGACATCCGAGCTTTGAGGCACTTTCTTCAAGTCCCAA 1022
Qy      1070  GATGAGGAGCTGTGGCGGCTTTTGTGATGACACCCAAAGGACATTTAGACGCTTACTA 1129
Db      1023  GATGAGGAGCTGTGGAGCGGCTTTTGTGATGAGCCCAAGGAGCATTTCAACAGCCCTACTA 1082
Qy      1130  TCCAGGCACTACCGGCGCAACATCTGACATGAGAAATCAAGGTGGCCCAACACG 1189
Db      1083  CCCAGGCACTACCGCACCAACATTTGACTGACATGAGAACATTTAGGTTGCCAACAA 1142
Qy      1190  GAACTGAAGTGTGGCTTCAAACCTTCTTATCTGTGTGAGACCCCAAGTACAGTGGCTC 1249
Db      1143  GATGTGAAGTGTGGCTTCAAATTTCTTCTACTGTGTGAGACCCGCGCTGTGCGGCGAC 1202
Qy      1250  CTGACCAAGGACTATGTGAGATCAACGGGGAGAACTACTCGGTGTGAGGTGCCAGTT 1309
Db      1203  CTGCCCCAAGGACTACTGTGAGATCAACGGGGAGAAATCTGCGGAGAGAGTCCAGTT 1262
Qy      1310  TGTGTGAGGAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTGTZAC 1369
Db      1263  GGTGTGACAGCAACAGCAAGATTACAGTTCGTTCCACTCACTCAATGACTCTTACAC 1322
Qy      1370  GGAACCGGCTTCTAGCTAGTACTCTCTTACAGTCTTCAACGACCCGTTGCCAGGAT 1429
Db      1323  CGACACCGGCTTCTTGTGAGAAATACCTCTCTACGACTCCAGTGAACCATGCGGGGCA 1382
Qy      1430  GTTCATGTGCAAGCTGGAACGCTGCACTCGAAGAGAACTGCGCTGCAAGGCTGGGCGA 1489
Db      1383  GTTCAGTGTGCGGACCGGGGAGTGTATCGAAGAGAGTGTGCTGTATGTGGGCGCA 1442
Qy      1490  CTGCCCCGATTATGATGATGAGGCTTACTGCGATGCAATGCCACCAAGTTCAGCTG 1549
Db      1443  CTGCAACGACCAACAGGATGAGCTCAACTGCAAGTTGGAAGCGGCGCACCAAGTTCA 1502
Qy      1550  CAAAAACAGTTCTGGAACCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGA 1609
Db      1503  CAAAAACAGTTCTGGAACCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGA 1562
Qy      1610  CCGAATGACAGAGAGGCTGACAGCTGTCTGTGGAAGTTTCAAGTTTCCAAATGGGA 1669
Db      1563  CAAAGGACAGAGAGGCTGAGTGTGTGCGGCCCAAGCTTCAAGTGTTCATATGGGA 1622
Qy      1670  GTGTCTCTCTTCAAGGCGAAGTGTATGTGAAGAGCAACTGTGAGATGGCTGTAGCA 1729
Db      1623  GTGCTCTGTGAAGAGCAGCAGTGTGAATGGGAAGAGCACTGTGGGAGCGGCTCGACA 1682
Qy      1730  GGCCTCATGTGACAGGTGAAATGTGTCTCTTGCACCAATATCTTACCGCTGCCAAA 1789
Db      1683  GGCCTCTGTGCGCAAGGTGAAGTGTGTCTGTACCAAAACACCTTACCGCTGTCTCAA 1742
Qy      1790  TGGCTCTGTGACCAAGGCGCAACCTGAGTGTATGGGAAAGAGAGCTGTAGCGATG 1849
Db      1743  TGGGCTCTGTGACCAAGGCGCAACCTGAGTGTATGGGAAAGAGAGCTGTAGCGAG 1802
Qy      1850  CTCGATGAGAAAACTGTGACTGTGGCTGTGATCTTTTACCAACAGGCTCGAGTGT 1909
Db      1803  CTCGATGAGAAAGACTGCGACTGTGGCTGTGATCTTTTACCAACAGGCTCGAGTGT 1862
Qy      1910  TGGTGGCAAGAAATGGCGACAGAGGCGAGTGGCTTGGGAGGTGAGCTTCAAGCCCTGG 1969
Db      1863  TGGGCGCAAGATGGGATGTGAGGGGAGTGGCTTGGGAGAGGAGCTTCAAGCCCTGG 1922
Qy      1970  CAGGCGCACTTGTGTGGGCTCGCTCATCTCTCTGACGAGGCTGTGCAAGCTCA 2029
Db      1923  CAGGCGCAACTGTGGGCTGTCTCTCATCTCTTCCAACTGGGCTGTGCGGCA 1982
Qy      2030  TTGCTTCAAGATGCAAAAAATTTCAAGTACTCAGACTCAACGATGTGAGCGGCTTCT 2089
Db      1983  CTGCTACATCATGACAGAGATTCAAGTACTCAAGCCCAAGCGAGTGAAGGTCTTCT 2042
Qy      2090  GGGTGTGCTGACCAAGAGCAAGCGAGTGTCTTGGGCTGAGAGCTGAAGCTCAAG 2149
Db      2043  GGGCTTGTGACCAAGAGCAAGCGAGGCTTGGGCTGAGAGGAGCGGAGCTCAAG 2102
Qy      2150  TATCATCAACCACTTCTCTCAATGATTTCACTTCACTTCACTATGACATGCGCTTGG 2209
Db      2103  CATCATCTCCACCTCTTCTTCAATGACTTCACTTCACTTCACTATGACATGCGCTTGG 2162
Qy      2210  GCTGAGAGAGTGGTGTGAGTACAGACCGTGTGTGGCCCACTTGCCTGTATGTATC 2269
Db      2163  GCTGAGAGAGCGGCGAGATGACCTCATGTGTGGCCCACTTGTGCTGTGGAGCGCTC 2222
```

QY 2270 CCATGTCCTTCCCTGCTGGAAGGCCATCTGGCTCAGAGCTGGGGGCAACAACAAAGAGG 2329  
DB 2223 CCATGTCCTTCCCTGCTGGAAGGCCATCTGGCTCAGAGCTGGGGGCAACAACCAATATG 2282  
QY 2330 AGGTACCGGAGGCGCTGATCTGTCAGAAAGGGTGGATCCGTGATCAACACAGACACTG 2389  
DB 2283 AGGCACTGGGCGCTGATCTGTCAGAAAGGGTGGATCCGTGATCAACACAGACACTG 2342  
QY 2390 TGAGACCTCATGCGCCAGACAGATCAACCCAGATGATGTGTGGGTTTCTCATGTG 2449  
DB 2343 CGAGAACCTCTGCGCGAGAGATCAAGCGCGCATGATGTGTGGGTTTCTCATGTG 2402  
QY 2450 GGGTGTGAGCTCTGCGAGAGGTGATCTGTGCGCCCTTGTCAAGCGCGGAGAAAGATG 2509  
DB 2403 CGCGGTGAGCTCTGCGAGAGGTGATCTGTGCGCGAGAGGTGATCTGTGCGCGGAG 2462  
QY 2510 GCGAATGTTCCAGGCTGTGTGTGAGCTGGGGTGAAGGCTGCGCTCAGAGAAACAAGCC 2569  
DB 2463 GCGGATCTTCCAGGCGCGGTGTGTGAGCTGGGGAGAGCGCTGCGCTCAGAGAAACAAGCC 2522  
QY 2570 AGGCGTGTACACAAGGCTCCCTGTAGTTGGGAGCTGGATCAAGAGACACTGGGGTATA 2629  
DB 2523 AGGCGTGTACACAAGGCTCCCTGTGTGGGAGCTGGATCAAGAGACACTGGGGTATA 2582  
QY 2630 GCGACATGACAGACAGCCCAACAACAACCCAGAGAGTCCCGCATGACACCTGG 2689  
DB 2583 GGGGCGCGGG---GCCACCCAAATGTGTACACTGCGGGGCGCAACCATGTCACCCAGT 2639  
QY 2690 ATACAGAGAGAGAAACATGACGACATTTATGCTGCTCCCGCCCGCAACAACAACCA 2749  
DB 2640 GTGTACG-CCTGCAAGCTGTGAGAGCTGAGACCGCTGACCTGCAACAGCGCC-CGAAACATA 2697  
QY 2750 GACTGTGACTGACATCTTATGAGACTCAGAGT 2780  
DB 2698 CACTGTGACTGACATCTTATGAGAGGCTCCAAAT 2728

RESULT 11  
AR081724 3147 bp DNA linear PAT 31-AUG-2000  
LOCUS AR081724  
DEFINITION Sequence 1 from patent US 5972616.  
ACCESSION AR081724  
VERSION AR081724.1 GI:10008450  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3147)  
AUTHORS O'Brien, T. J. and Tanimoto, H.  
TITLE TAG-15: an extracellular serine protease overexpressed in breast  
JOURNAL Patent: US 5972616-A 1 26-OCT-1999;  
FEATURES  
source 1..3147  
Location/Qualifiers  
/organism="Unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 61.6%; Score 1883.2; DB 6; Length 3147;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATGGAACCGCCAAACCATGAGTACGATCGGGGCGGAGAGGGGGCTTCAG 104  
DB 5 GAGCGGCTCGGGGTAACATGAGGAGCGATCGGGCGCGCAAGGGGAGGGGGCCGAAG 64  
QY 105 GACTTCGGGCGGAGCTCAAGTACACTCCGGCTTAGAGACATGATGCTTTGAGAG 164  
DB 65 GACTTCGGGCGGAGCTCAAGTACACTCCGGCTTAGAGACATGATGCTTTGAGAG 124  
QY 165 GGATGTGAGATCTCTGCTCGAACAATGCGCAAGAAAGTGAAGAGGAGCGCCCAAGCGC 224

DB 125 GCGGTGAGATTCCTGCACTCAACAACTCAAGAAAGTGAAGAAAGCATGGCCCGGCGCG 184  
QY 225 TGGAGTGTCTGATGAGCACTGCTGATGCTCTTGTCTCTGCTCAATGCTGGCTTG 284  
DB 185 TGGAGTGTCTGATGAGCACTGCTGATGCTCTTGTCTCTGCTCAATGCTGGCTTG 244  
QY 285 CTGTGTGCACTTCATTAAGGAAATGTCGGGTTTCAAAAAGTCTTCAATGGCCATG 344  
DB 245 CTGTGTGCACTTCATTAAGGAAATGTCGGGTTTCAAAAAGTCTTCAATGGCCATG 304  
QY 345 AGGATCAACAAATGAAATCTTCTGATGCTGATGAGAACTCCCACTCCACAGATTATC 404  
DB 305 AGGATCAACAAATGAAATCTTCTGATGCTGATGAGAACTCCCACTCCACAGATTATC 364  
QY 405 AGCCTGAGCAGCAGGATGAAAGAGGCGCTGAAAGCTGCTATCAATGAATGCTGTCTG 464  
DB 365 AGCCTGAGCAGCAGGATGAAAGAGGCGCTGAAAGCTGCTATCAATGAATGCTGTCTG 424  
QY 465 GGTCTTACCAAGAAAGTGGCTGTATGCTCTTCAATGAGGAGCTGATCGGCTAC 524  
DB 425 GCGCCCTTACCAAGAAAGTGGCTGTATGCTCTTCAATGAGGAGCTGATCGGCTAC 484  
QY 525 TACTGTGCAAGTTCAGATCCCGCCCACTGAGCAAGAAAGTGTATCGGCTATGCT 584  
DB 485 TACTGTGCAAGTTCAGATCCCGCCCACTGAGCAAGAAAGTGTATCGGCTATGCT 544  
QY 585 GTGAGGCAAGTGTAACTATGCAACCCCGAGCAGGCGCATGAATCTTGTGCTTACA 644  
DB 545 GAGGAGGCGGTATGATGATGCTGCGCCCGCGGCGCGCTCTGAATGCTTGTGTGCTACC 604  
QY 645 TGTGTGTGCTTCCCTGATGACCCCAAGATGCTGAGAGAGCTGAGCAACAGCTGCG 704  
DB 605 TCAGTGTGCTTCCCTGATGACCCCAAGATGCTGAGAGAGCTGAGCAACAGCTGCG 664  
QY 705 AGTTTTCCTGCAATGCGCAATGAGGAGAGCTGACAGCTTCACTACCCCTGCTTCCC 764  
DB 665 AGTTTTCCTGCAATGCGCAATGAGGAGAGCTGAGAGCTTCACTACCCCTGCTTCCC 724  
QY 765 AACAGTCCCTTACCCCGGAGATGCGCGCTGCAAGTGTGCTGCGGGGAGCGGCACTCT 824  
DB 725 GACAGCCCTTACCCCGGAGATGCGCGCTGCAAGTGTGCTGCGGGGAGCGGCACTCT 784  
QY 825 GTGTGAGCTCACTTCCGAAGTGTGATGCTGCTGCTGATGAGCATGAGGAGTAC 884  
DB 785 GTGTGAGCTCACTTCCGAAGTGTGATGCTGCTGCTGATGAGCATGAGGAGTAC 844  
QY 885 CTGTGACCGGTATGATGCTGAGCTGAGCCCAAGAAACCCAGCTGTGTGCTGTGT 944  
DB 845 CTGTGACCGGTATGATGCTGAGCTGAGCCCAAGAAACCCAGCTGTGTGCTGTGT 904  
QY 945 GGCACCTTTCACCCCTCTTACACCTGACTTCTCTCTCTCCAGAACGTCTTCTTGTCTC 1004  
DB 905 GGCACCTTTCACCCCTCTTACACCTGACTTCTCTCTCTCCAGAACGTCTTCTTGTCTC 964  
QY 1005 AGCTGATTAACCAATATCTGACCGGAGCATCTGCTTGAAGGCACTTCTTCAGCTG 1064  
DB 965 ACATGATTAACCAATATCTGACCGGAGCATCTGCTTGAAGGCACTTCTTCAGCTG 1024  
QY 1065 CCCAAGATGAGAGCTGTGCGGCTTTTGAATGACACCCCAAGGAGCATTTAGACGCTCC 1124  
DB 1025 CCTAGATGAGAGCTGTGAGAGCGGCTTATGATTAAGGCGGAGCATTTCAACAGCTCC 1084  
QY 1125 TACTATCCAGGCACTTACCCCGCCCAATCACTGACATGAGAAATATCAAGTGTCCCAAC 1184  
DB 1185 AACCGAACTGAAAGTGTGCTTCAACTCTTCTTATGAGTGAAGCCCAACTACAGTGTG 1244  
QY 1145 AACCGAACTGAAAGTGTGCTTCAACTCTTCTTATGAGTGAAGCCCAACTACAGTGTG 1204  
DB 1245 GGTCTCTGCAACAGGATATGAGATCAACGGGAGAAATGATCTGCGGTGAGAGTCTC 1304  
QY 1205 GGTCTCTGCAACAGGATATGAGATCAACGGGAGAAATGATCTGCGGTGAGAGTCTC 1264





QY	345	AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTCCACTCCACAGAGTTTATC	404
Db	305	AGGATCAAAATGAGATTTTGTGATGCTACGAAACTCCAACTCACTGATTTGTA	364
QY	405	AGCTGGCCAGCGAGGTGAAGGAGGGCGCTGAAGCTCTGTACATGAATGATCCCTCTCTG	464
Db	365	AGCTGGCCAGCAAGGTGAAGGAGCGGCTGAAGCTCTGTACAGCGGAATCCCATTTCTG	424
QY	465	GGTCCCTACCAAGAAGTGGCTGTAACTGCCTTCAGTGAGGGCAGTCAATCGCTAC	524
Db	425	GGGCCCTACCAAGAAGTGGCTGTGAGCGGCTTCAAGGAGGGCAGCTCATCGCTAC	484
QY	525	TACTGTCAAGATTCAAGCATCCCCCACAACCTTGGCAGAAAGGTTATGGCCAAATGGCT	584
Db	485	TACTGGTCTGAGTTCAAGCATCCCGCAGCAACTGTGTGAGAGGGCCAGAGCGCTCATGGCC	544
QY	585	GTGAGAGGAGTGTGAACATTGCCACCCCGAGCACTGGAATCTTCTGTCTTACA	644
Db	545	GAGGAGGCGGTAGTCATGCTGCCCCCGGGGCGGCTCCTGGAATCTTTGTGGTCAAC	604
QY	645	TCTGTGGTGGCTTCCCATTTGACCCCGAATGCTGCAGAGACTCAAGACAAACAGCTGC	704
Db	605	TCACTGGTGGCTTTCCCAACGGAATCCAAACAGTACAGAGACCCAGAGCAACAAGCTGC	664
QY	705	AGTTTGGCCTGCAATGCCATGCTGTCAGACAGTGCACAGCTTACTACTCCCTGGCTTCCCC	764
Db	665	AGCTTTGGCCTGCAAGCCCGGGGTGTGAAGCTGATGGCTTCAACAGCCCGGCTTCCCT	724
QY	765	AACAGTCCCTACCGGCGGATGCCCGCTGCAGTGGGTCTCTGGGGGGGAGCGCGCACTCT	824
Db	725	GACAGCCCTACCCCGCTCATGCCCGCTGCGCAGTGGGCGCTTGGGGGGAGCGCGCACTCA	784
QY	825	GTGCTGAGGCTCACCTTCCGAAGCTTTGATGTGCTCTCTGTATGAGCATGCGAATGAC	884
Db	785	GTGCTGAGGCTCACCTTCCGCAAGTTTGACTTGTCTGTGCAAGCGCGCAGAGCAGAC	844
QY	885	CTGTGCACCGGTATGATAGCTTGAGGCCCACTGGAACCCCAACGCTGTGGTGGCGGTGT	944
Db	845	CTGTGTGACGGGTATCAACACCTTGAAGCCCACTGAGCGCCCAAGCTGTGGTGAAGTGTGT	904
QY	945	GSCACTTCTCAACCCCTCTACAACTGACTTCTCTCTCTCCAGAAAGCTTCTCTGTGC	1004
Db	905	GSCACTTCAACCTCTCTCTTAACAAGCTTCACTCTCTCCAGAAAGTCTGTGATC	964
QY	1005	ACGCTGATTAACCAATACGACCGGCGCAATCTGGCTTTGAGGCCACTTCTTCAAGCTG	1064
Db	965	ACACTGATTAACCAACACTGAGGGGGGCGATCCCGGCTTTGAGGCCACTTCTTCAAGCTG	1024
QY	1065	CCCAAGATGACGACTGTGGCGGCTTTTGTAGTGACACCCCAAGGACATTTAGAGGCCCC	1124
Db	1025	CCTAGAGATGAGACGCTGTGGAGGCGGCTTACGTAAAGCCAGGAGACCTTCAACAGGCCCC	1084
QY	1125	TACTATCCAGGCACTAACCCCGCCCAACTCAATGCAATGGAATATATCAAGGTGCCCAAC	1184
Db	1085	TACTATCCAGGCACTAACCCCAACTGACTGCAATGGAACATTTGAAGTCCCAAC	1144
QY	1185	AACCGGAAGTGAAGGTGGCTTCAAACTTTATCTGTGTGACCCCAAGTACAGTG	1244
Db	1145	AACCAAGACGTGAAGGTGAGCTTCAAAATTTCTTCACTGTGTGAGCCCGGCTGCTGGC	1204
QY	1245	GGCTCTGCAACCAAGACTATGTGAGATCAACGGGAGAAAGTACTGGGGTGAAGAGTCC	1304
Db	1205	GGCACTGCGCCCAAGACTACGTGAGATCAATGGGAGAAATATGTGGAGAGAGGTCC	1264
QY	1305	CAGTTTGTGTAGAGCAAGCAAGCAAGCAAGATTACAGTCACTTCTGATCACTCG	1364
Db	1265	CAGTTTGTGTACCAAGCAACAGCAAGCAAGATCAAGTTTCTTCACTCAAGATTAAGTCC	1324
QY	1365	TACACGAGACAGGGGTTCTTAGTGAATACCTTCTTACAGTCTTCCACAGACCCGCTGCCA	1424
Db	1325	TACACGAGACAGGGGTTCTTAGTGAATACCTTCTTACAGTCTTCCAGTGAACCATGCGCG	1384

QY	1425	GGGATGTTCA	TGTGCAAG	CTGAGCGGTGCAT	TCGAAAGAA	CTGCGCTGCGACGGCTGG	1484
Db	1385	GGGGAGTTTCA	CGTGCACG	ACGGGGCGGTGTAT	CCGGAAAGAA	CTGCGCTGTATGATGCGTGG	1444
QY	1485	GCAGATCTGCCG	GATTTATAGTA	TGAGCGGTAT	ACTGCGCAT	TGGAATGGCA	CCCA
Db	1445	GCCGATCTGCAC	CCGACCA	CACGCGATGAGCT	CACTGCA	GTGGAGCGCGGCGCACCA	GTTC
QY	1545	ACGNGCAAAA	ACAGTCTT	CGCAAGCCCT	CTTCTGGGT	CTGTGACA	GTGTCA
Db	1505	ACGTGCAAGAA	CAAGTTCT	CTGCAAGCCCT	CTTCTGGGT	CTGTGACA	GTGTGAA
QY	1605	GGGACGGA	AGTGA	CGAGAGGCGCT	GCAGTGTCT	CTGCTGGG	AGTTTCAA
Db	1665	GGAGACA	AAACAGCA	CGACGAGCGGTGCAG	TTGTTC	CCGGCCAA	GCCTTCA
QY	1725	GACAGGCTT	CATGTG	ACAGCGTGAAT	GTGCTCT	TTGACA	CAAAATATA
Db	1785	CAAATG	GCCTCT	GTGTGACAA	AGGCGCA	CCCTGAGT	GTGAT
QY	1845	GATGGCTT	CCCATG	AGAAAA	CTGTGA	CTGTGGGCT	GCATCTTT
Db	1905	GACGCTCA	GTAGAA	AGACTGTGCAT	CTGTGGGCT	GCATCTTT	ACAAACAGCT
QY	1965	CTGGCGC	AGGCGCA	CTTGTGTGGG	CCTCGCTCA	CTCTCTCT	GTACGTGGCT
Db	2025	GCTCAT	TGCTTT	CAGAGTAC	CAAAAA	TTTCAAGT	ACTACACTA
QY	2085	TTCTCGG	GTCTGCTG	AGCAAGCA	AGCGCAT	CTCTGGG	GTGCAGAG
Db	2145	AAACG	TATCAT	CA	CCACCTT	CCTTCA	TATTTCA
QY	2205	CTGAG	AGCTG	AGAAAT	CGGTG	AGTACAGCA	CCGCGT
Db	2265	GCTAC	CAATGT	CTTCCCTG	CTTGC	AGGCA	CTTGTGGGT
QY	2325	GAAGG	AGTAC	CGAGCGT	GTATCT	CTGACAA	AGGAT
Db	2385	TAAGG	AGTAC	CGAGCGT	GTATCT	CTGACAA	AGGAT
QY	2445	AGTGG	AGTGT	GGACTCT	CGACAGG	GTGTCT	GTGACG
Db	2505	GATGG	AGTGT	GGACTCT	CGACAGG	GTGTCT	GTGACG

Db 2465 GATGGGCGGATCTTCCAGGCGGCTGGTGGTGGTGGGAGCGGCTGGCTCAAGAGAAC 2524  
 QY 2565 AAGCCAGGCGTGTACACAAAGGCTCCCTGTAGTTCG9GACTGGATCAAAAGACACATCGG 2624  
 Db 2525 AAGCCAGGCGTGTACACAAAGGCTCCCTGTAGTTCG9GACTGGATCAAAAGAGAACATCGG 2584  
 QY 2625 GTATAGAGATGAGACAGACAGCCGACCAAAACCCACAGAGGATGCCGACATGACACA 2684  
 Db 2585 GTATAGGCGCGCGG--GCCACCCCAATGTGTACACCTGGCGGCGCCACCATCGTCCACC 2641  
 QY 2685 CCGTATACAGAGAGAGAACATGACACATTTATGCTGGGCGCTCCCGCCCAACACA 2744  
 Db 2642 CCAAGTGTGACG--CTGCAAGCTGTGAGACTGACCGCTGACTGACACAGCGCC--CCAGA 2699  
 QY 2745 ACCCAGACTGTGAATGAGATCTTTAGGACTCAGAGT 2780  
 Db 2700 ACATACACTGTGAATCTCAATCTCCAGGGGCTCCAAAT 2735

RESULT 13  
 AR229712/c AR229712 3147 bp RNA linear PAT 20-DEC-2002  
 LOCUS AR229712 Sequence 18 from patent US 6451500.  
 DEFINITION AR229712  
 ACCESSION AR229712  
 VERSION AR229712.1 GI:27269442  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 3147)  
 AUTHORS Leon J.W.  
 TITLE Imaging member containing heat switchable carboxylate polymer and method of use  
 JOURNAL Patent: US 6451500-A 18 17-SEP-2002;  
 FEATURES Location/Qualifiers  
 source 1..3147  
 /organism="Unknown"  
 /mol\_type="unassigned RNA"

## ORIGIN

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;

Best Local Similarity 81.2%; Pred. No. 0;

Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATCGGACCGCAAAACCATGGGTAGCAATCGGGCGCGCAAGCGCGAGGGGCTCTCAG 104  
 Db 3143 GAGCGGCTCGGGGTACATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGCCGGAAG 3084  
 QY 105 GACTTCGGGGCGGGACTCAAGTCAACTCCCGCTAGAGACATGAATGGCTTTGAGAG 164  
 Db 3083 GACTTCGGGGCGGGACTCAAGTCAACTCCCGCGACAGAAAGTGAATGGCTTGAGAGAA 3024  
 QY 165 GGTGTGAGATTCTGTGCTCGAAACATGCGCAAGAAAGTGAAGAGAGAGGCGCCAGGCGC 224  
 Db 3023 GCGGTGAGATTCTGTGCTCGAAACATGCGCAAGAAAGTGAAGAGAGAGGCGCCAGGCGC 2964  
 QY 225 TGGGTGTGCTGTGGGCAAGTGTGTTCACTTCTTGTCTCTCCCTCAATGAGCTGGCTTG 284  
 Db 2963 TGGGTGTGCTGTGGGCAAGTGTGTTCACTTCTTGTCTCTCCCTCAATGAGCTGGCTTG 2904  
 QY 285 CTGGTGTGCACTTCCATTTATGGAATGTGCGGGTTCAAAAAGTCTTCAATGAGCTCATG 344  
 Db 2903 CTGGTGTGCACTTCCATTTATGGAATGTGCGGGAGTGTTCATGAGCTCATG 2844  
 QY 345 AGGATCACAAAATGAGATCTTTCTGATGCGTATGGAAGTCCAGCTCCACAGAGTTTATC 404  
 Db 2843 AGGATCACAAAATGAGATTTTGTGATGCTAGGAACTCCAACTCCAGAGTTTATA 2784  
 QY 405 AGCTGGCGCAGCAGTGAAGAGCGGTGAAGTGTGCTGTATCAATGAAGTCCCTGTCTG 464  
 Db 2783 AGCTGGCGCAGCAGTGAAGAGCGGTGAAGTGTGCTGTATCAAGTGTCCATTTCTG 2724

QY 465 GGTCTTACCAAGAAAGTGGCTGTATCTGACCTTCAAGTGAAGGCAAGTGTATCGCTAC 524  
 Db 2723 GGCCCTTACCAAGAAAGTGGCTGTATCTGACCTTCAAGTGAAGGCAAGTGTATCGCTAC 2664  
 QY 525 TACTGTCAAGATTCAAGATTCGCCACACACTGGCAGAAAGTTGATCGCCATGCT 584  
 Db 2663 TACTGTCAAGATTCAAGATTCGCCACACACTGGTGAAGAGGCGGAGCGGTATGCTC 2604  
 QY 585 GTGAGCGAATTGTAATTTGCAATTCGCAACCCCGAGCAAGGCGACTGTAAATCTTGTCTTACA 644  
 Db 2603 GAGAGGCGGTATGATGCTGCCCCCGGGCGGCTCCCTGAAGTCTTGTGTATACC 2544  
 QY 645 TCTGTGTGCTTCCCATTTGACCCAGATGCTGAGAGAGTCAAGCAACAGCTGC 704  
 Db 2543 TCAGTGTGCTTTCCTCCAGAGCTCCAAACAGTACAGAGAGCCAGAGACAGCTGC 2484  
 QY 705 AGTTTGGCTGTATGCTGACAGAGTGAACAGCTTCACTACCTCGCTGCTCC 764  
 Db 2483 AGCTTGGCTGTATGCTGACAGAGTGTGAGCTGATGAGCTTCAACAGCGCTTCCCT 2424  
 QY 765 AACAGTCCCTACCGCGGAGATGCGCGCTGCAATGGGTCTGTGGGGGAGCGCTACT 824  
 Db 2423 GACAGCCCTTACCCGCTATGCTGCTGCAAGTGGCCCTGTGGGGGAGCGCTACTCA 2364  
 QY 825 GTGCTGAGCTCACTTCCGAAAGCTTGAATGTGCTCCCTGTGATGAGCATGAGCACTGAC 884  
 Db 2363 GTGCTGAGCTCACTTCCGAGCTTGAACCTTGTGCTGCTGCAAGAGCGCGGAGGAC 2304  
 QY 885 CTGCTACCGTGTATATAGCTGAGAGCCCAATGAAACCCAGCTGTGTGCGGCTGT 944  
 Db 2303 CTGCTACCGTGTATATAGCTGAGAGCCCAATGAAACCCAGCTGTGTGCTGTGT 2244  
 QY 945 GGCACCTTCAACCCGCTTCAACACTGATCTTCTCTCCAGAAAGCTTCTTGTTC 1004  
 Db 2243 GGCACCTTCAACCCGCTTCAACACTGATCTTCTCTCCAGAAAGCTTCTTGTTC 2184  
 QY 1005 AGCGTATTAACCAATATGACCGGAGACATCTGTGCTTGAAGGCACTTCTTCCAGCTG 1064  
 Db 2183 ACACGTATTAACCAATATGACCGGAGACATCTGTGCTTGAAGGCACTTCTTCCAGCTG 2124  
 QY 1065 CCCAAGATAGAGCTGTGGCGGCTTTTGAAGTGAACCCCAAGGACATTTAGCAGCCCC 1124  
 Db 2123 CCTAGATAGAGCTGTGGCGGCTTGAAGTGAACCCCAAGGACATTTAGCAGCCCC 2064  
 QY 1125 TACTATCCAGGCACTACCGCGGCCAATCACTGACATGGAATATCAAGTGTCCCAAC 1184  
 Db 2063 TACTATCCAGGCACTACCGCGGCCAATCACTGACATGGAATATCAAGTGTCCCAAC 2004  
 QY 1185 AACCGAAAGTGAAGTGTGCTTCAAACTCTTCTATGTGTGAGACCCCAAGTATCCAGTG 1244  
 Db 2003 AACCGAAAGTGAAGTGTGCTTCAAACTCTTCTATGTGTGAGACCCCAAGTATCCAGTG 1944  
 QY 1245 GGTCTCTGACCAAGAGACTATGTGAGATCAACGGGAGAAAGTACTGCGGTGAAGTCTC 1304  
 Db 1943 GGTCTCTGACCAAGAGACTATGTGAGATCAACGGGAGAAAGTACTGCGGTGAAGTCTC 1884  
 QY 1305 CAGTTGTGTGAGAGAGCAACAGAGCAAGATTAAGTCCACTTCAATTTATCACTG 1364  
 Db 1883 CAGTTGTGTGAGAGAGCAACAGAGCAAGATTAAGTCCACTTCAATTTATCACTG 1824  
 QY 1365 TACACGAGACCGGCTTCTAGCTGAGTACCTCTCTACGACTCCACGAGCCGTCGCA 1424  
 Db 1823 TACACGAGACCGGCTTCTAGCTGAGTACCTCTCTCTACGACTCCACGAGCCGTCGCA 1764  
 QY 1425 GGGATGTTCATGTGAAGAATGAGCGGTGATCCGAAAGAACTGCGGTGCGAGCGCTG 1484  
 Db 1763 GGGATGTTCATGTGAAGAATGAGCGGTGATCCGAAAGAACTGCGGTGCGAGCGCTG 1704  
 QY 1485 GCAAGCTGCGCGGATTAAGAGTGAAGGTTAAGTCCAGTGAATGCCAACCAAGTTC 1544  
 Db 1703 GCAAGCTGCGCGGATTAAGAGTGAAGGTTAAGTCCAGTGAATGCCAACCAAGTTC 1644  
 QY 1545 AGGTGCAAAAACCAAGTTCGCAAGCCCTTCTGTGGTGTGTGACAGTGTCAAGACTGT 1604





RESULT 15  
 AR430938/c  
 LOCUS AR430938 3147 bp RNA 1linear PAT 18-DEC-2003  
 DEFINITION Sequence 18 from patent US 6649741.  
 ACCESSION AR430938  
 VERSION AR430938.1 GI:40192680  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1. (bases 1 to 3147)  
 O'Brien, T.J. and Taniuchi, H.  
 TADG-15: an extracellular serine protease overexpressed in  
 carcinomas  
 JOURNAL Patent: US 6649741-A 18 NOV-2003;  
 FEATURES  
 source Location/Qualifiers  
 1..3147  
 /organism="unknown"  
 /mol\_type="unassigned RNA"  
 ORIGIN  
 Query Match 60.6%; Score 1883.2; DB 6; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

Db 2483 AGCTTGGCCCTGACGCGCCGCGGTGTGAGCTGATGCGCTTACCAAGCCCGGCTTCCCT 2424  
 QY 765 AACAGTCCCTAACCCGCGCGATGCGCGCTCCAGTGGCTCTGCGGGGGAGACGCGCATCTC 824  
 Db 2423 GACAGCCCTTACCCCGCTTATGCGCTGCGAGTGGGCCCTGGGGGGGAGCGCGCATCTCA 2364  
 QY 825 GTGCTGAGCCTCACTTCCGAAAGCTTGTATGTCCTCTGTATGATGATGATGATGATGATG 884  
 Db 2363 GTGCTGAGCCTCACTTCCGAGCTTGTGACCTTGCCTCCGACAGCGCGCGAGCGAGC 2304  
 QY 885 CTGCTGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944  
 Db 2303 CTGCTGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2244  
 QY 945 GGCACCTTCTACCCCTCTCAACACTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004  
 Db 2243 GGCACCTTCT 2184  
 QY 1005 AGCGTATTAACCAATCTGACCGCGCATCTGCGCTTGTAGGCTTGTAGGCTTGTAGGCT 1064  
 Db 2183 ACACTGATTAACCAACTGACCGCGCATCTGCGCTTGTAGGCTTGTAGGCTTGTAGGCT 2124  
 QY 1065 CCGAATGATGACGCTGTGCGGCTTTTGTAGTACACCCAGGACATTTAGAGCGCC 1124  
 Db 2123 CTTAGATGACGCTGTGAGCGCTTCTGTAAGCCCAAGGAGCATTTCAAGCGCC 2064  
 QY 1125 TACTATCAGAGCCACTACCCGCCCAACATGACATGACATGAAATATCAAGTGTCCCAAC 1184  
 Db 2063 TACTATCAGAGCCACTACCCGCCCAACATGACATGACATGAAATATCAAGTGTCCCAAC 2004  
 QY 1185 AACCGAAGCTGAAAGTGTGCTTCAACTCTTATCTGTGTGACCCCAAGTACAGCTG 1244  
 Db 2003 AACCGAAGCTGAAAGTGTGCTTCAACTCTTATCTGTGTGACCCCAAGTACAGCTG 1944  
 QY 1245 GGCCTCTGCAACAAAGCTATGTGAGATCAACGGGAGAAATGACTGCGTGAAGGTTC 1304  
 Db 1943 GGCACCTGCCCCAAGGACTGATGAGATCAATGGGAGAAATGACTGCGGAGAGGTTC 1884  
 QY 1305 CAGTTTGTGTGAGCAGCAACAGCAGCAAGATTAACGTCACTTCCATTTGTATCACTCG 1364  
 Db 1883 CAGTTTGTGTGAGCAGCAACAGCAGCAAGATTAACGTCACTTCCATTTGTATCACTCG 1824  
 QY 1365 TACACGACACCGGCTTCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1424  
 Db 1823 TACACGACACCGGCTTCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1764  
 QY 1425 GGGATGTTCAATGTGCAAGACTGACGCGTCAATCGAAAGAACTGCGTGCAGCGCTGG 1484  
 Db 1763 GGGATGTTCAATGTGCAAGACTGACGCGTCAATCGAAAGAACTGCGTGCAGCGCTGG 1704  
 QY 1485 GCAGACTGCCCCGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1544  
 Db 1703 GCGGACTGCAACGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1644  
 QY 1545 ACCTGCAAAAACCACTTCTGGAAGCCCTCTTGGGCTGTGACAGTGTCAACAGACTGT 1604  
 Db 1643 ACCTGCAAAAACCACTTCTGGAAGCCCTCTTGGGCTGTGACAGTGTCAACAGACTGT 1584  
 QY 1605 GGGGACGAAATGACGAAAGGCTGTGACGCTGTCTGTGTGGAGTTTCAAGTGTTCAT 1664  
 Db 1583 GGGGACGAAATGACGAAAGGCTGTGACGCTGTCTGTGTGGAGTTTCAAGTGTTCAT 1524  
 QY 1665 GGGAGTGTCTCTCTCAGAGCCAGAGTGTATGGAAGAGACAACTGTGTGAGATGAGGTCT 1724  
 Db 1523 GGGAGTGTCTCTCTCAGAGCCAGAGTGTATGGAAGAGAGAGATGTGTGGAGAGCGGTCTC 1464  
 QY 1725 GACGAGCTTCAATGTGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1784  
 Db 1463 GACGAGCTTCTCTCTCAGAGGTGAGAGGTGATGATGATGATGATGATGATGATGATGATG 1404  
 QY 1785 CAAAATGAGCTGTGTCTGAGCAAGGCAACCTTGAATGTGTGTGTGTGTGTGTGTGTGTGT 1844  
 Db 1403 CTCAATGGCTCTGCTTGTGAGCAAGGCAACCTTGAATGTGTGTGTGTGTGTGTGTGTGTGT 1344



QY 1845 GATGGCTCCGATGAGAAAACCTGTACTGTGGGCTCGATCCTTTACCAACAGGCTCGC 1904  
 Db 1343 GACGGCTCAATGAGAGAGACTGTGGCTGTGGCTGATTCACAGAGACAGGCTCGT 1284  
 QY 1905 GTGGTTGTGGACGAATGCGAGACGAGGGAGTGGCCCTGGGAGGTGAGCTTCCAGGCC 1964  
 Db 1283 GTTCTTGGGGGACGAGATGGGATGAGGGCGAGTGGCCCTGGGAGGTGAGCTTCCAGGCC 1224  
 QY 1965 CTGGGCCAGGGCCACTTGTGTGGGGCTCGCTCATCTCTCTGACTGGCTGTCTGTGA 2024  
 Db 1223 CTGGGCCAGGGCCACTTGTGTGGGGCTCGCTCATCTCTCTGACTGGCTGTCTGTGA 1164  
 QY 2025 GCTCATTTGCTTTAGAGATGACAAAATTTCAAGTACTCAGACTACAGATGTGAGCGGCC 2084  
 Db 1163 GCACACTGTCAATCAGATGACAGAGATTCAGGTACTCAGACCCCAAGCAGTGGAGCGGCC 1104  
 QY 2085 TTCTGGGCTGTGTGGACGAGACGAGCGGAGGCTGTGGGGTGGAGAGCTGAAGCTC 2144  
 Db 1103 TTCTGGGCTGTGTGGACGAGACGAGCGGAGGCTGTGGGGTGGAGAGCTGAAGCTC 1044  
 QY 2145 AAACGTATCATACCCACCTTCTTCAATGATTTCACTTCGACTATGACATCGCCTTG 2204  
 Db 1043 AAGCGCATCATCTCCACCTTCTTCAATGATTTCACTTCGACTATGACATCGCCTTG 984  
 QY 2205 CTGGAGCTGAGAGAGTGGTGGATGACAGACCGTGTGGCGCCCATTTGCTGCTGAT 2264  
 Db 983 CTGGAGCTGAGAGAGTGGTGGATGACAGACCGTGTGGCGCCCATTTGCTGCTGAT 924  
 QY 2265 GCTACCATGTCTTCTCTGTGGACGAGCGCATTTGGGTCAAGGCTGGGGGACACAAA 2324  
 Db 923 GCTTCCCATGTCTTCTCTGTGGACGAGCGCATTTGGGTCAAGGCTGGGGGACACAAA 864  
 QY 2325 GAGGAGGTACCGAGCGGCTGATCTCTGAGAAAGGTGAGATCGGTCAATCAACAGACC 2384  
 Db 863 TATGAGGCACTGAGCGGCTGATCTCTGAGAAAGGTGAGATCGGTCAATCAACAGACC 804  
 QY 2385 ACCTGTGAGAGACTCATGCGCGACAGATCAACCCACGATGATGTGTGGGTTTCTTC 2444  
 Db 803 ACCTGTGAGAGACTCATGCGCGACAGATCAACCCACGATGATGTGTGGGTTTCTTC 744  
 QY 2445 AGTGGGGGTGTGGAATCTCTGACAGGGTGACTGTGGGGCTTGTCAAGCGCGAGAAA 2504  
 Db 743 AGCGGCGGCTGTGGAATCTCTGACAGGGTGACTGTGGGGCTTGTCAAGCGCGAGAAA 684  
 QY 2505 GATGGGCGAATGTTCAAGCTGTGTGTGAGCTGGGTGAAGCTGCGCTCAGAGGAAC 2564  
 Db 683 GATGGGCGAATGTTCAAGCTGTGTGTGAGCTGGGTGAAGCTGCGCTCAGAGGAAC 624  
 QY 2565 AAGCGAGGCTGTGATACAAAGGCTTCTGTGATGTGGGACTGGATCAAAAGACACTGGG 2624  
 Db 623 AAGCGAGGCTGTGATACAAAGGCTTCTGTGATGTGGGACTGGATCAAAAGACACTGGG 564  
 QY 2625 GTATAGCAGATGACAGACAGCGGACCAAAACCAACGAGGGATGCCGACATGCA 2684  
 Db 563 GTATAGGAGCGCGG--GCCACCCAAATGTGATACCTGCGGGGCCACCCATGTCCACC 507  
 QY 2685 CTTGATACAGAGAGAACTGACGATTTATGTGTGGCTTCCCCCCCCCAACACA 2744  
 Db 506 CCAAGTGTACAG--CTTGGAGCTGGAGCTGGACCGCTGACTGACACAGCGGCC--CCAGA 449  
 QY 2745 ACCCAGACTGTGAATGATCTCTTGTGACTCAGAGT 2780  
 Db 448 ACATTAAGTGTGAATCTCTTGTGACTCAGAGT 413

BLANK  
PAGE